



# LUND UNIVERSITY

## Automatic segmentation in CMR - Development and validation of algorithms for left ventricular function, myocardium at risk and myocardial infarction

Tufvesson, Jane

2015

[Link to publication](#)

*Citation for published version (APA):*

Tufvesson, J. (2015). *Automatic segmentation in CMR - Development and validation of algorithms for left ventricular function, myocardium at risk and myocardial infarction*. [Doctoral Thesis (compilation), Clinical Physiology (Lund)]. Lund University (Media-Tryck).

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

## About the author



Jane Tufvesson, born Sjögren, has early on had an interest in mathematics and a goal set on a master in engineering. With a specialization in medical image analysis the road from the master thesis to this PhD thesis seems as a straight line. During the PhD studies in the Cardiac MR group, Jane has also pursued a career as a developer, quality assurance manager and vice president in the company Medviso AB and has been on maternity leave with her two kids, now two and three years old. During her maternity leave Jane revived her interest in sewing and designing clothes and has

aimed for having all me-made clothes on her PhD defense. Along with the sewing and PhD studies Jane has also pursued her interest in personal development and leadership by becoming an accredited coach.



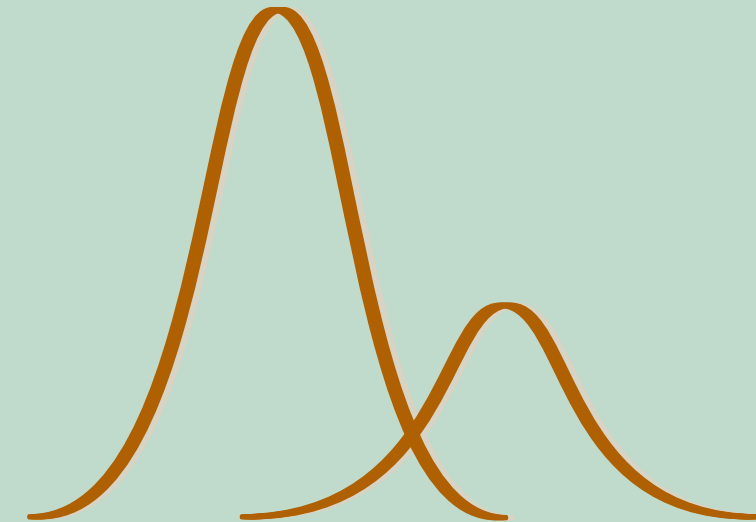
Lund University  
Department of Biomedical Engineering  
Faculty of Engineering  
ISBN 978-91-7623-508-9  
ISRN LUTEDX/TEEM--1098--SE



# Automatic segmentation in CMR

Development and validation of algorithms for left ventricular function, myocardium at risk and myocardial infarction

JANE TUFVESSON



Lund University  
Faculty of Engineering  
Department of Biomedical Engineering



# Automatic segmentation in CMR

Development and validation of algorithms for  
left ventricular function, myocardium at risk  
and myocardial infarction

Jane Tufvesson



**LUND**  
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Engineering, Lund University, Sweden.  
To be defended in Lecture Hall GK Biomedical Center, Lund University  
on Friday December 4th at 9:00.

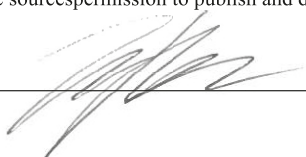
*Faculty opponent*

Professor Dara L. Kraitchman, Johns Hopkins Medicine,  
Baltimore, Maryland, USA

Organization LUND UNIVERSITY		Document name
Author(s) Jane Tufvesson		Date of disputation
		Sponsoring organization
Title and subtitle Automatic Segmentation in CMR; Development and validation of algorithms for left ventricular function, myocardium at risk and myocardial infarction		
<p>Abstract</p> <p>In this thesis four new algorithms are presented for automatic segmentation in cardiovascular magnetic resonance (CMR); automatic segmentation of the left ventricle, myocardial infarction, and myocardium at risk in two different image types. All four algorithms were implemented in freely available software for image analysis and were validated against reference delineations with a low bias and high regional agreement.</p> <p>CMR is the most accurate and reproducible method for assessment of left ventricular mass and volumes and reference standard for assessment of myocardial infarction. CMR is also validated against single photon emission computed tomography (SPECT) for assessment of myocardium at risk up to one week after acute myocardial infarction. However, the clinical standard for quantification of left ventricular mass and volumes is manual delineation which has been shown to have a large bias between observers from different sites and for myocardium at risk and myocardial infarction there is no clinical standard due to varying results shown for the previously suggested threshold methods.</p> <p>The new automatic algorithms were all based on intensity classification by Expectation Maximization (EM) and incorporation of <i>a priori</i> information specific for each application. Validation was performed in large cohorts of patients with regards to bias in clinical parameters and regional agreement as Dice Similarity Coefficient (DSC). Further, images with reference delineation of the left ventricle were made available for future benchmarking of left ventricular segmentation, and the new automatic algorithms for segmentation of myocardium at risk and myocardial infarction were directly compared to the previously suggested intensity threshold methods.</p> <p>Combining intensity classification by EM with <i>a priori</i> information as in the new automatic algorithms was shown superior to previous methods and specifically to the previously suggested threshold methods for myocardium at risk and myocardial infarction. Added value of using <i>a priori</i> information and intensity correction was shown significant measured by DSC even though not significant for bias. For the previously suggested methods of infarct quantification a poorer result was found in the new multi-center, multi-vendor patient data than in the original validation in animal studies or single center patient studies. Thus, the results in this thesis also show the importance of using both bias and DSC for validation and performing validation in images of representative quality as in multi-center, multi-vendor patient studies.</p>		
Key words		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		Language: English
ISSN and key title		ISBN 978-91-7623-508-9
Recipient's notes	Number of pages 143	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature



Date

20151027



# Automatic segmentation in CMR

Development and validation of algorithms for  
left ventricular function, myocardium at risk  
and myocardial infarction

Jane Tufvesson



**LUND**  
UNIVERSITY

Copyright © Jane Tufvesson

Faculty of Engineering  
Biomedical Engineering

Doctoral Thesis in Biomedical Engineering  
ISRN LUTEDX/TEEM – 1098 – SE  
Report 2:2015

ISBN 978-91-7623-508-9 (print) 978-91-7623-509-6 (pdf)

Printed in Sweden by Media-Tryck, Lund University  
Lund 2015



To my greatest source of inspiration  
Erik, Jon and Ellen.

*People take different roads to seeking fulfillment and happiness.  
Just because they're not on your road doesn't mean they've gotten lost.*  
-Dalai Lama



# Preface

Pursuing a PhD has been a journey in itself and during this time my life has changed tremendously and I have developed personally. I would like to extend my gratitude to all those who have been by my side during this journey and special thanks to:

My main supervisor **Einar Heiberg** for always believing in me and giving me so many opportunities to grow. **Håkan Arheden** for the constant inspiration towards personal development. My co-supervisor **Henrik Engblom** for all our conversations which have given me so much energy and happiness. My co-supervisor **Marcus Carlsson** for giving me the extra inspiration to becoming a coach.

To those who have become my friends during the journey. **Katarina Steding-Ehrenborg** for always being there and helping me reflect on who I want to be. **Ulrika Pahlm-Webb** for the fun time in Orlando and the continued support in the following hardships. **Erik Hedström** for the wonderful conversations and inviting me to the world of musicals and classy bars in London. **Mikael Kanski** for walking alongside me during the whole journey. **Helen Fransson** for always being there and for inviting me to the trip where I met my husband. **Robert Jablonowski** for being an inspiration and support especially in the end of the PhD studies and for your equal fondness of structure.

To all my co-workers in the cardiac MR group, the Department of Clinical Physiology and Medviso AB for being the workplace which I have loved going to for so many years. For all the fun in the chamber and at conferences **Johannes Töger**, **Per Arvidsson**, **David Nordlund**, **Tom Gyllenhammar**, **Mariam Al-Mashat**, **Shanaz Ahkil** and **Felicia Seeman**. **Giorgios Kantasis** and **Sebastian Bidhult** for your enthusiastic view on MRI. **Nils Lundahl** for standing tall of Medviso. **Jakob Almer**, **Petter Frieberg**, **Fredrik Hedeer** and **Sverrir Stephensen** for pursuing the journey as PhD students at the same time as other work commitments. **Joey Ubachs**, **David Strauss** and **Christos Xanthis** for pursuing a PhD with us here in Lund. **Pia Sjöberg** and **Karin Markenroth-Bloch** for being amazing roommates at conferences. To the post-docs **Ellen Ostenfeldt**, **Magnus Hansson**, **Cecilia Hindorf**, **Jenny Oddstig**, **Fredrik Testud** and **Jelena Bock** for your continued commitment to science. To **Bo Hedén**, **Henrik Mosén**, **Ann-Helen Arvidsson**, **Christel Carlander**, **Lotta Åkesson** and **Reza Farazdaghi** for your invaluable efforts and connection to the clinic. **Johanna Koul**, **Nina Aaviden** and **Tone Eikrem Harr** for all the fun we have had outside work. **Kerstin Brauer** and **Karin Larsson** for all your administrative help. Last but not least **Anthony Aletras** and **Jonas Jögi** for being inspirational leaders.

To those who welcomed me at the department of Biomedical Engineering. **Johan Nilsson** and **Désirée Jarnebrandt** for helping me in the transfer. **Fredrik Ejserholm**, **Tobias Erlöv**, **Belinda Adler** and **Hanifeh Khayyeri** for being familiar face at the new department.

To my parents **Janni Sjögren** and **Janne Sjögren**, my sister **Jannike Sjögren** and my brother **Henrik Lundh** for always believing in me and never being surprised by the new challenges I have taken on.

Finally the most important part of my journey, my beloved family. My husband **Erik Tufvesson**, for always supporting me, and our wonderful children **Jon** and **Ellen**, you mean the world to me.

The work in this thesis has been supported by the Swedish Research Council, the Knowledge foundation, Swedish Heart and Lung Foundation, Region of Scania, Medical Faculty at Lund University and the European Commission 7th framework programme.

# Contents

List of papers	1
Author's contributions	2
Summary	3
Populärvetenskaplig sammanfattning	5
Abbreviations	7
Introduction	9
Aim	11
Background	13
The human heart	13
Cardiac pumping	13
Myocardial infarction	15
Magnetic resonance imaging	17
Assessment of cardiac pumping	18
Infarct imaging	18
Edema imaging	20
Image segmentation	22
Intensity classification	22
Segmentation methods	23
Validation	24
Methods	25
Automatic segmentation algorithms	25
<i>A priori</i> information	27
Intensity classification by Expectation Maximization	30
Validation	31
Software for cardiac image analysis (Paper I)	31
Left ventricular segmentation (Paper II)	31
Myocardium at risk in T2-weighted CMR (Paper III)	31
Myocardium at risk in CE-SSFP (Paper IV)	32
Myocardial infarction (Paper V)	33

Results and comments	35
Software for cardiac image analysis (Paper I)	35
Left ventricular segmentation (Paper II)	35
Myocardium at risk (Paper III and IV)	38
Myocardial infarction (Paper V)	43
Conclusions	49
Bibliography	51



# List of papers

In this thesis the following papers are presented:

- I. Heiberg E, Sjögren J, Ugander M, Carlsson M, Engblom H, Arheden H; *Design and validation of Segment - freely available software for cardiovascular image analysis*. BMC Medical Imaging, 10:1, 2010.
- II. Tufvesson J, Hedström E, Steding-Ehrenborg K, Carlsson M, Arheden H, Heiberg E; Validation and development of a new automatic algorithm for time-resolved segmentation of the left ventricle in magnetic resonance imaging. BioMed Research International, 970357, 2015.
- III. Sjögren J, Ubachs JFA, Engblom H, Carlsson M, Arheden H, Heiberg H; *Semi-automatic segmentation of myocardium at risk in T2-weighted cardiovascular magnetic resonance*. J Cardiovasc Magn Reson, 12(1) p25, 2010.
- IV. Tufvesson J, Carlsson M, Aletras AH, Engblom H, Deux JF, Koul S, Sörensson P, Pernow J, Atar D, Erlinge D, Arheden H, Heiberg H; *Automatic segmentation of myocardium at risk from contrast enhanced SSFP CMR: validation against expert readers and SPECT*. Submitted.
- V. Tufvesson J<sup>\*</sup>, Engblom H<sup>\*</sup>, Jablonowski R, Carlsson M, Aletras AH, Hoffmann P, Jacquier A, Kober F, Metzler B, Erlinge D, Atar D, Arheden H, Heiberg E; *A new automatic algorithm for quantification of myocardial infarction in late gadolinium enhancement cardiac magnetic resonance images: experimental validation and implementation in multi-center, multi-vendor patient data*. Manuscript.

<sup>\*</sup>The authors contributed equally to the study

## Author's contributions

This section describes my contributions to each of the papers.

- I. I implemented large parts of the general object segmentation module, designed and implemented the test script and revised the manuscript.
- II. I contributed to concept and design of the study, developed and implemented the automatic segmentation algorithm, analyzed and interpreted results, and drafted and revised the manuscript.
- III. I contributed to concept and design of the study, developed and implemented the automatic algorithm, implemented the threshold methods, analyzed and interpreted segmentation results, and drafted and revised the manuscript.
- IV. I helped conceive the study, contributed to concept and design of the study, developed and implemented the automatic algorithm, analyzed and interpreted segmentation results, and drafted and revised the manuscript.
- V. I conceived the study, contributed to concept and design of the study, developed and implemented the automatic algorithm, analyzed and interpreted segmentation results, and drafted and revised the manuscript.

# Summary

In this thesis four new algorithms are presented for automatic segmentation in cardiovascular magnetic resonance (CMR); automatic segmentation of the left ventricle, myocardial infarction, and myocardium at risk in two different image types. All four algorithms were implemented in freely available software for image analysis and were validated against reference delineations with a low bias and high regional agreement.

CMR is the most accurate and reproducible method for assessment of left ventricular mass and volumes and reference standard for assessment of myocardial infarction. CMR is also validated against single photon emission computed tomography (SPECT) for assessment of myocardium at risk up to one week after acute myocardial infarction. However, the clinical standard for quantification of left ventricular mass and volumes is manual delineation which has been shown to have a large bias between observers from different sites and for myocardium at risk and myocardial infarction there is no clinical standard due to varying results shown for the previously suggested threshold methods.

The new automatic algorithms were all based on intensity classification by Expectation Maximization (EM) and incorporation of *a priori* information specific for each application. Validation was performed in large cohorts of patients with regards to bias in clinical parameters and regional agreement as Dice Similarity Coefficient (DSC). Further, images with reference delineation of the left ventricle were made available for future benchmarking of left ventricular segmentation, and the new automatic algorithms for segmentation of myocardium at risk and myocardial infarction were directly compared to the previously suggested intensity threshold methods.

Combining intensity classification by EM with *a priori* information as in the new automatic algorithms was shown superior to previous methods and specifically to the previously suggested threshold methods for myocardium at risk and myocardial infarction. Added value of using *a priori* information and intensity correction was shown significant measured by DSC even though not significant for bias. For the previously suggested methods of infarct quantification a poorer result was found in the new multi-center, multi-vendor patient data than in the original validation in animal studies or single center patient studies. Thus, the results in this thesis also show the importance of using both bias and DSC for validation and performing validation in images of representative quality as in multi-center, multi-vendor patient studies.



# Populärvetenskaplig sammanfattning

I denna avhandling presenteras fyra nya metoder för att automatiskt mäta hjärtats förmåga att pumpa ut blod i kroppen, mäta storleken på en hjärtinfarkt samt mäta hur stor del av hjärtmuskeln som riskerade att dö vid hjärtinfarkten. Dessa nya metoder kan användas för att minska skillnader mellan mätningar gjorda av olika läkare, på olika sjukhus och i olika forskargrupper. Detta gör mätresultaten mer tillförlitliga och kan i sin tur leda till att diagnostik och behandling av hjärt-kärlsjukdomar förbättras.

Vid en hjärtinfarkt har en blodpropp bildats i ett av de små blodkärlen som försörjer hjärtmuskeln med blod och en del av hjärtmuskeln drabbas då av syrebrist. Om inte syrebristen åtgärdas snabbt riskerar den delen av hjärtmuskeln att dö. För att patienten även efter hjärtinfarkten ska få en god livskvalitet är det viktigt att rädda så mycket som möjligt av hjärtmuskeln och därmed bevara hjärtats funktion. Därför är det viktigt att få bra behandling i tid. För att utvärdera nya behandlingar kan man beräkna hur mycket av hjärtmuskeln som räddades genom att mäta både storleken på det område av hjärtmuskeln som riskerade att dö och det slutgiltiga området som dog. Detta kräver dock tillförlitliga mätmetoder.

Med nya förbättrade behandlingar kan infarktens storlek minimeras vilket leder till en ökad livskvalité för patienten och mindre risk för följsjukdomar som till exempel hjärtsvikt. När patienten drabbas av hjärtsvikt innebär det att hjärtat inte har förmåga att pumpa ut så mycket syresatt blod som kroppen behöver. I takt med att hjärtat får ett allt tyngre arbete försvagas hjärtat ytterligare och en ond spiral bildas. Hjärtsvikt kan oftast inte botas men patientens livskvalité kan förbättras med diagnos och behandling. För diagnos och behandling av hjärt-kärlsjukdomar är det viktigt att tillförlitligt mäta både hjärtats pumpförmåga och storlek.

Med bilder från magnetkamera kan man noggrant mäta hjärtats pumpförmåga, hjärtinfarktens storlek och riskområdets storlek. Mätningarna görs genom att rita i bilderna och detta görs ofta manuellt av läkarna. Det skiljer dock mellan mätningar gjorda av en erfaren och en oerfaren läkare samt att det skiljer mellan erfarna läkare på olika sjukhus. För att åtgärda detta har flera automatiska metoder utvecklats men trots att manuell utlinjering har stora begränsningar har ingen tidigare automatisk metod visats tillräckligt bra för att ersätta manuell utlinjering. För att nya automatiska metoder ska kunna användas som standard på både sjukhus och i forskning behöver metoderna utvärderas i många patienter och bilder med olika bildkvalité. Dessutom behöver andra forskare än de som utvecklat den automatiska metoden ha möjlighet att utvärdera de nya metoderna. Därför är datorprogrammet och de nya automatiska metoderna som vi utvecklar fria att användas inom forskning.

I **delarbete I** presenteras det datorprogram där de nya automatiska metoderna har implementerats. I **delarbete II** presenteras en ny automatisk metod för att mäta hjärtats storlek och pumpförmåga. I **delarbete III och IV** presenteras två olika automatiska metoder för att mäta riskområdet vid en hjärtinfarkt i två olika typer av bilder. I **delarbete V** presenteras en ny metod för att mäta hjärtinfarktens storlek. Gemensamt för de fyra nya metoderna är att de utvecklades i nära samarbete med läkarna som annars gör manuella utlinjeringar. Grunden för att kunna utveckla metoder som automatiskt ritas i bilderna från en magnetkamera är att kunna skilja på mörka och ljusa områden i bilderna. Bilderna som de fyra metoderna utvecklades för har olika bildkvalité men ändå kunde samma metod användas för att särskilja mörka och ljusa områden. Denna metod kallas EM-algoritmen och den visades vara bättre på att särskilja mellan mörkt och ljust än tidigare metoder. De fyra nya metoderna har i studie II-V utvärderats i patienter och jämförts med manuella utlinjeringar utförda av erfarna läkare. Utvärdering gjordes både genom att mäta skillnad i storlek men också genom att mäta hur pass bra områdena som ritades automatiskt stämde överens med de manuellt utlinjerade.

De fyra automatiska metoderna gav resultat som stämde väl överens med manuellt utlinjerade områden, både vad gäller storlek och överlapp mellan områdena. Skillnaden mellan automatisk och manuell mätning av hjärtats storlek, riskområdet och hjärtinfarktens storlek var nästan lika liten som mellan två erfarna läkare inom samma forskargrupp. De nya metoderna kan användas för att minska skillnaden mellan mätningar utförda av olika läkare, framförallt på olika sjukhus eller med olika erfarenhet. Att använda EM-algoritmen för att särskilja mellan ljusa och mörka områden tillsammans med expertkunskaper från läkarna visades vara bättre än tidigare utvecklade metoder. Dessutom visades vikten av att göra utvärderingar med både skillnad i volymer och överlapp mellan område och att göra denna utvärdering i många patienter från olika sjukhus.

# Abbreviations

AHA: American Heart Association  
bSSFP: balanced steady state free precession  
CE-SSFP: contrasted enhanced steady state free precession  
CMR: cardiovascular magnetic resonance  
DSC: Dice similarity coefficient  
ECG: electrocardiogram  
EDV: end-diastolic volume  
EF: ejection fraction  
ESV: end-systolic volume  
EM: expectation maximization  
FWHM: full width half maximum  
IR: inversion recovery  
LAD: left anterior descending coronary artery  
LCx: left circumflex coronary artery  
LGE: late gadolinium enhancement  
LM: left main coronary artery  
LVM: left ventricular mass  
%LVM: percentage of left ventricular mass  
MaR: myocardium at risk  
MI: Myocardial infarction  
MRI: magnetic resonance imaging  
MVO: microvascular obstruction  
P2C: point to curve distance  
PCI: percutaneous coronary intervention  
PSIR: phase sensitive inversion recovery  
RCA: right coronary artery  
RF: radio frequency  
SD: standard deviations  
SPECT: single photon emission computed tomography  
STEMI: st-elevation myocardial infarction  
SV: stroke volume  
T2-STIR: T2 weighted triple inversion turbo spin echo  
TI: inversion time  
T1: spin-lattice property of tissue  
T2: spin-spin property of tissue  
TTC: triphenyltetrazolium chloride





# Introduction

Cardiovascular magnetic resonance (CMR) is the most accurate and reproducible method for quantifying left ventricular mass and volumes [1]. It is reference standard for assessment of myocardial infarct size [2-3] and has been validated against SPECT for assessment of myocardium at risk after acute myocardial infarction [4-5]. For quantification it is necessary to delineate left ventricular myocardial borders, and hyper-enhanced regions of the myocardium representative of myocardial infarction and myocardium at risk. Manual delineation is clinical standard for left ventricular mass and volumes and has high reproducibility for experienced observers within single centers but has recently been shown to have a large bias between experienced observers from different centers [6]. For quantification of myocardial infarct size and myocardium at risk both manual delineation, simple threshold methods and advanced automatic segmentation methods have been suggested but there is no accepted clinical standard [7].

To reduce observer dependency and time for analysis automatic segmentation methods are desirable. For a new method to be considered clinical standard it needs to be validated and available for clinical use. However, automatic segmentation algorithms tend to become complex and thereby hard to re-implement and many promising algorithms never reach clinical routine if not implemented in already existing software.

Automatic quantification of the left ventricle is commonly implemented in software for CMR analysis and there is a rich literature on the multiple methods for automatic quantification of the left ventricle [8-9]. However, there is a continued use of manual delineation which indicates that previously suggested methods do not give satisfactory results. Not all segmentation methods have included the most basal part of the left ventricle in which the segmentation is complicated by the long axis motion of left ventricle causing the outflow tract to move in and out of the imaging plane during a heartbeat. The long axis motion is a major contributor to the stroke volume [10] and previous automatic segmentation methods which have included the outflow tract have either not defined the detection of outflow tract [11] or long axis motion [12], or has detected the long axis motion manually [13]. Further it is hard to conclude on the superiority of the suggested methods since the results have been obtained on images from different patient populations and with different image quality [9].

For quantification of myocardium at risk in T2-weighted CMR several methods have been suggested such as manual delineation [4] and thresholds by two standard deviations from remote (2SD) [14-15], full width half maximum (FWHM) [16] and Otsu's threshold [17] and have been evaluated for intra- and inter-observer variability and test-retest scans [18]. More advanced automatic segmentation algorithms with regional analysis [19-20] have also been suggested but not compared head-to-head with the threshold methods. Threshold methods was shown less reproducible than manual delineation when taking into account intra- and inter-observer variability and test-retest scans [18]. For manual delineation observers take into account both regional intensity differences, transmuralities and *a priori* knowledge on perfusion territories. This is not taken into account by the threshold methods nor was the perfusion territories taken into account in two proposed automatic segmentation algorithms [19-20]. Assessment of MaR by contrast enhanced SSFP (CE-SSFP) is not as established as T2-weighted imaging and only manual delineation [5] has

been proposed for quantification of MaR. Although both T2-weighted imaging and CE-SSFP have been used in two clinical trials and CE-SSFP was shown to produce more diagnostic images [21].

For quantification of myocardial infarct (MI) size in late gadolinium enhancement (LGE) multiple methods have been proposed. Manual delineation, thresholds by  $n$ -SD from remote, FWHM [22] and Otsu [23], and more advanced automatic algorithms [24-25] have been validated in experimental studies [22, 25-26] and patient studies [26-28] with varying results. Manual delineation was also for MI quantification shown more reproducible than threshold method when considering test-retest scans [18] in addition to inter- and intra-observer variability [29]. Further, clinical assessment of myocardial infarct size is performed using either of two techniques, inversion recovery (IR) and phase sensitive inversion recovery (PSIR) and a bias has been shown between the threshold methods in IR and PSIR images [30]. The more advanced automatic algorithms [24-25] included regional analysis and accounted for partial volume effects but have only been developed and validated for either IR or PSIR images in single center studies.

# Aim

The overall aim was to develop and supply the research community in cardiovascular magnetic resonance with freely available, well validated algorithms for automatic segmentation of left ventricular function, myocardium at risk and myocardial infarction.

The specific aims for each study were:

- I. Present the design and validation of the cardiovascular image analysis software Segment and to announce its release in a source code format.
- II. Develop an automatic algorithm for time resolved segmentation of the left ventricle from the basal slices with outflow tract to the apex, validate the algorithm with regard to clinical parameters and regional agreement for comparison to previous methods and to enable benchmarking for future algorithms.
- III. Develop and validate an automatic algorithm for segmentation of myocardium at risk in T2-weighted images which uses *a priori* knowledge on the appearance of MaR and cardiac anatomy.
- IV. Develop the automatic segmentation of myocardium at risk for contrast enhanced SSFP images and validate the new algorithm against manual delineation in multi-center, multi-vendor patient data and against SPECT.
- V. Develop an automatic algorithm for segmentation of myocardial infarction which can account for both partial volume effects and varying image quality as seen in magnitude inversion recovery and phase sensitive inversion recovery LGE images and to validate the new algorithm experimentally and in multi-center, multi-vendor patient data.



# Background

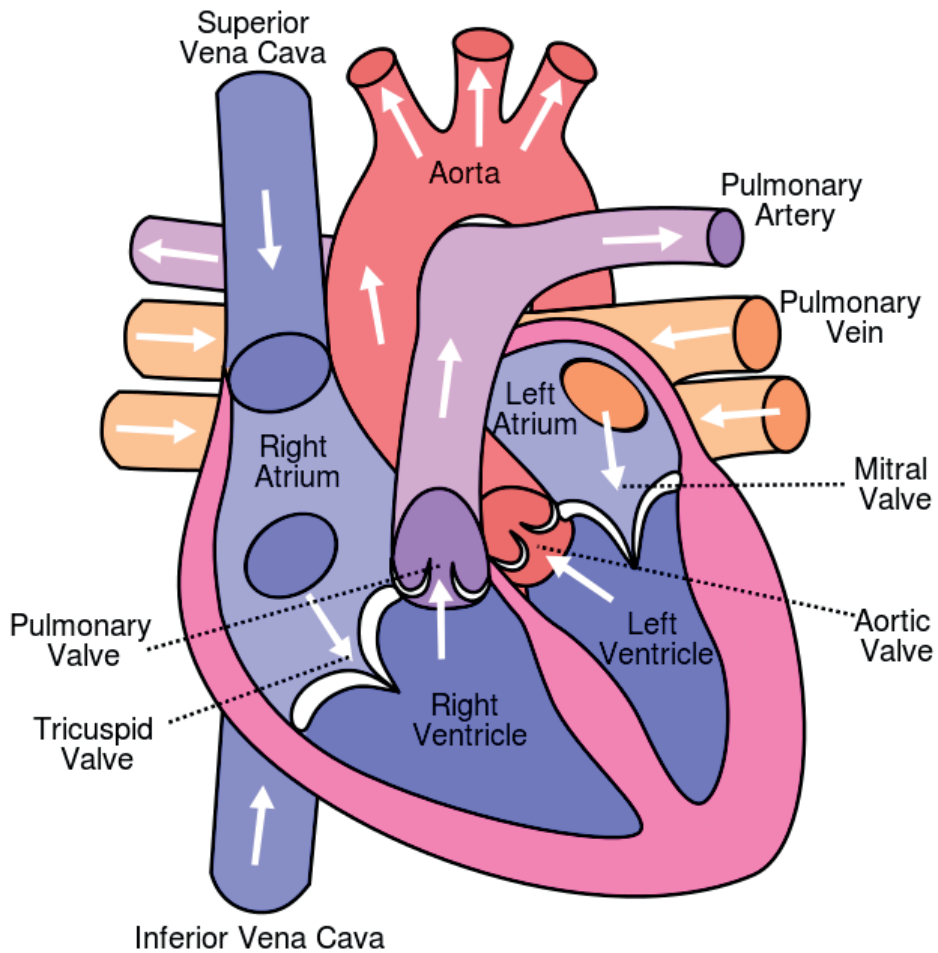
## The human heart

The heart is a muscle which has one important task; to supply the whole body with oxygenated blood. The heart consists of four chambers, one atrium and one ventricle on the right and left side. The atria and ventricles are separated by the atrio-ventricular plane in which four valves are situated, through which the blood flows when open. Blood flows from the right atria (RA) through the tricuspid valve into the right ventricle (RV) from which it is pumped through the pulmonary valve into the pulmonary artery. The blood is oxygenated in the lungs and returned to the left side of the heart via the pulmonary veins to the left atria (LA) and flows through the mitral valve into the left ventricle (LV). The left ventricle pumps the blood through the aortic valve into the aorta and further out in the body to supply muscles and organs with oxygenated blood. Deoxygenated blood is returned to the heart via superior and inferior vena cava. The heart muscle is called myocardium and is thicker in LV than RV due to the higher pressure in the systemic circulation. **Figure 1** shows a schematic drawing of the blood flow through the human heart.

## Cardiac pumping

Cardiac pumping can be divided into two phases, contraction or ejection, called systole, and relaxation or filling, called diastole. During systole the myocardium of LV and RV contracts and blood is ejected to the aorta and pulmonary artery and during diastole the myocardium relaxes and blood from the atria is filling the ventricles. The ejected blood volume is called stroke volume (SV) and is the difference between the blood volume after filling, end-diastolic volume (EDV), and the blood volume after ejection, end-systolic volume (ESV). The stroke volume is also expressed as percentage of the end diastolic volume and then called ejection fraction (EF).

The contraction of the myocardium causes the atrio-ventricular valve plane to move towards the base, the myocardium to thicken, and the myocardium to move inwards. The combined effect of the thickening and the inward motion of the myocardium cause the inner border of the myocardium, the endocardium, to move inwards and the outer border of the myocardium, the epicardium, to move slightly inwards or outwards. The atrio-ventricular plane displacement, also called long axis motion, causes the outflow tract in the



**Figure 1 Overview of the blood flow through the human heart**

Schematic drawing of the blood flow through the human heart. Right atrium - right ventricle - pulmonary artery - lungs - pulmonary veins - left atria - left ventricle - aorta - body - inferior and superior vena cava.

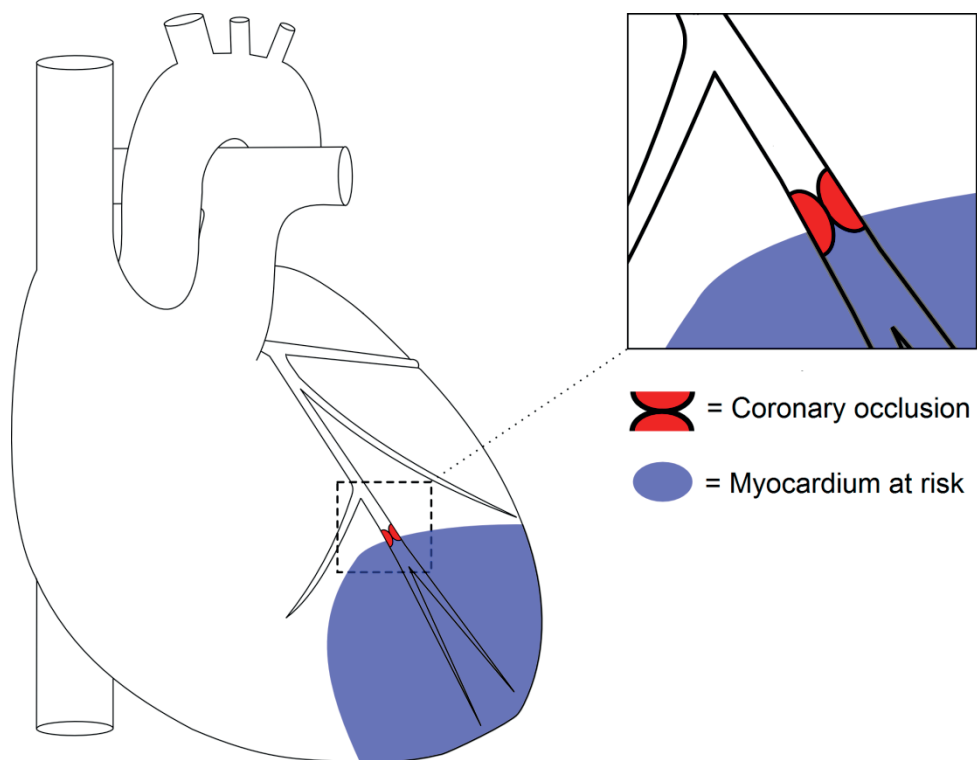
*Image used with permission from author (Creative Commons Attribution-ShareAlike (CC BY-SA) 3.0): <http://en.wikipedia.org/wiki/User:Wapcaplet>*

most basal slices to move towards the apex in end-systole and is a major contributor to the stroke volume [10].

For analysis of left ventricular volumes and mass the papillary muscles should be included in the blood volume and excluded from the myocardial mass [7]. The papillary muscles are situated in the ventricles and connected to the mitral and tricuspid valve and contracts as the myocardium contracts. In end-systole it is hard to separate the endocardial border from the papillary muscles but it is important to pay attention to the separation between the papillary muscles and the endocardial border during the heart beat since both the myocardial and papillary muscle mass is constant over time.

## Myocardial infarction

In order for the heart to contract normally, the myocardium itself also needs blood supply through the coronary arteries. The coronary arteries originate from the root of the aorta as right coronary artery (RCA) and the left main artery (LM), which is divided into the left anterior descending artery (LAD) and left circumflex artery (LCx). The coronary arteries are responsible for the blood supply to different perfusion territories and if the blood supply is insufficient ischemia occurs in the affected region. The ischemic region is called myocardium at risk (MaR) in the acute setting of an occluded coronary artery and the ischemic region is then at risk for myocardial cell death, necrosis, also called myocardial infarction (MI). **Figure 2** shows a schematic view of a coronary occlusion LAD with subsequent MaR. To save as much myocardium as possible it is important to restore perfusion as soon as possible, which can be done pharmacologically and by invasive methods of percutaneous coronary intervention (PCI) or coronary artery bypass surgery. The amount of myocardial salvage depends on the time to reperfusion as the myocardial infarction propagates from the endocardial border against the epicardial border in a wavefront manner [31] to finally become a transmural MI, extending the full MaR region, if blood flow is not restored (**Figure 3**). Reperfusion of the ischemic region can however also cause injuries to the myocardium and new treatments are constantly under development to reduce myocardial injury and improve treatment of myocardial infarction. One such myocardial injury is microvascular obstruction (MVO) [32] which is a region of no blood flow within the previously ischemic region. Efficacy of new treatments can be assessed as the infarct size alone or be put in relation MaR as myocardial salvage index (MSI),  $MSI = 1 - MI/MaR$ . By assessing efficacy as MSI instead of only MI the number of patients needed in clinical trials can be reduced and new treatments can be developed faster. Both MI [2-3] and MaR [33] can be assessed by CMR days, up to one week, after treatment of MI and thus efficacy of treatment can be evaluated in one single CMR examination.

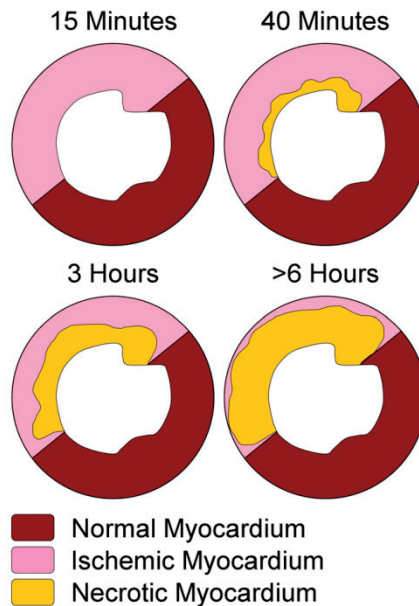


**Figure 2 Schematic representation of a heart with coronary occlusion**

A coronary occlusion, shown in red, causes a region of no blood flow, represented in blue, which becomes ischemic myocardium at risk (MaR), that may become necrotic if the occlusion is not resolved and thus becoming myocardial infarction.

*Images used and modified with permission from the author Jablonowski R: Assessment of myocardial viability using magnetic resonance imaging .Lund University 2015.*





**Figure 3 Schematic representation of the wavefront propagation of myocardial infarction**

A coronary occlusion in the left anterior descending artery is represented in a short-axis slice of the left ventricle. With increasing duration of ischemia, myocardial necrosis (yellow) will evolve within the ischemic myocardium (pink) in a wavefront manner from the endocardial border to the epicardial border. Myocardial salvage can be seen as the remaining part of the ischemic myocardium, myocardium at risk, which was not turned into necrotic myocardium, myocardial infarction.

*Image used with permission from the author Ubachs J: Quantitative and qualitative assessment of the myocardium at risk. Lund University 2011.*

## Magnetic resonance imaging

Magnetic resonance imaging (MRI) is an imaging modality free from ionization which can obtain images of the body in any imaging plane and which can be designed to acquire contrast between a variety of different tissues and organs. MRI exploits the magnetic properties of the hydrogen atoms, which for example are present in water and fat in our bodies. There are four main components to the MR scanner, firstly, a strong magnetic field called  $B_0$  of typically 1.5 or 3T, with which the hydrogen atoms are either aligned to or against once the patient is placed in the scanner. The hydrogen atoms also spins around its own nuclei with a specific frequency, called the Larmor frequency. This Larmor frequency is exploited by the second main part of the MR scanner, the radio frequency (RF) transmitter. The RF transmitter modifies, or flips, the main magnetic field  $B_0$  by transmitting RF pulses proportional to the Larmor frequency and the  $B_0$  field to create resonance. An electric signal is created which can be detected by the third main part of the MR scanner, a receiver. The final main part of the MR scanner is gradient coils which affect the magnetic field in order to encode spatial information to the signals and thus an MR image can be created.

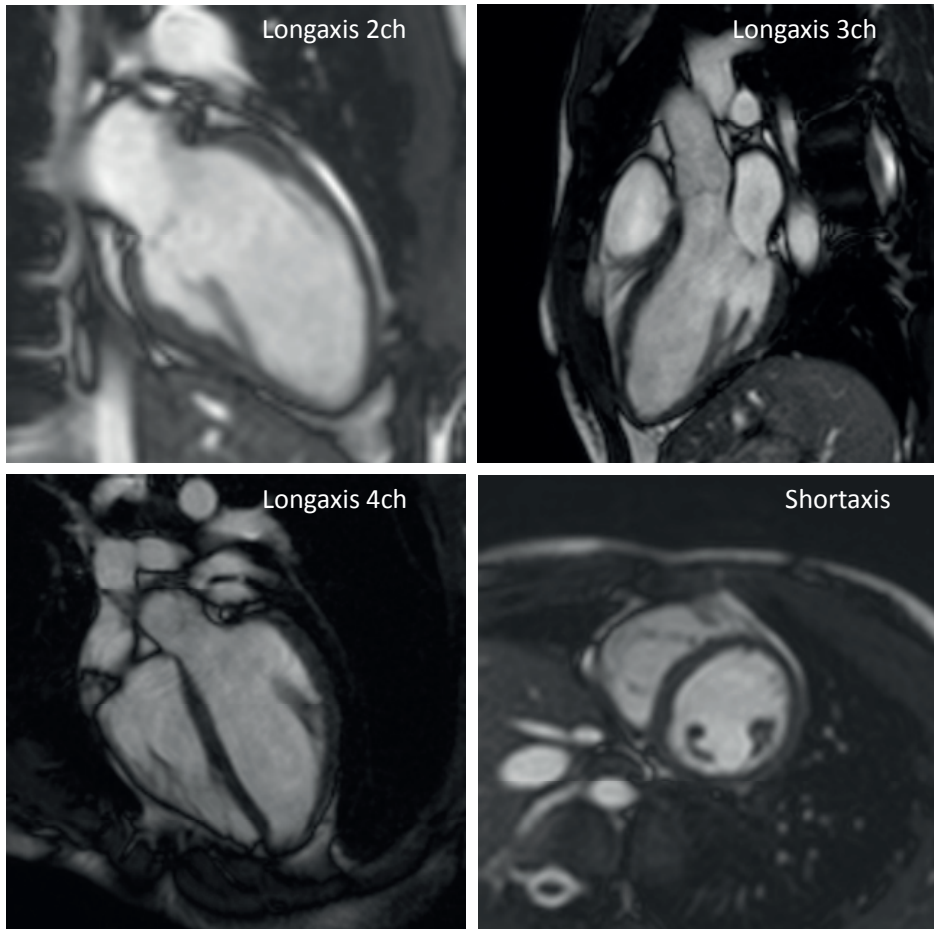
The signal created in the MR scanner is affected by two magnetic properties of the tissue, T1 and T2 relaxation, which both can be exploited to create contrast between different tissues. T1 relaxation is due to the energy from the RF pulse, being transmitted to neighboring molecules which causes a decay of the signal over time. T2 relaxation is due to the spins affecting each other, called spin-spin interaction, which also causes a decay of the signal, which is faster than the T1 relaxation. By programming the MR scanner with different RF pulse sequences the T1 and T2 properties can be exploited to create contrast between different types of tissue. For cardiovascular magnetic resonance (CMR) contrast is for example created between blood and myocardium, normal and pathological myocardium, such as edema and myocardial infarction.

## Assessment of cardiac pumping

Assessment of cardiac function requires imaging with good temporal resolution and a good contrast between blood and myocardium. Balanced steady state free precession (bSSFP) is based on the ratio between T2 and T1 and the RF pulses can be applied with short intervals, repetition times, and thus can acquire both good contrast and temporal resolution. However, there are two basic challenges in imaging the heart, the heart is beating and the heart is moving due to the breathing of the patient. To overcome these difficulties imaging is synchronized to the beating of the heart by using an electrocardiogram (ECG) and either the breathing of the patient can be detected and compensated for or the imaging can be performed during breath hold. Magnetic resonance imaging can be performed in any imaging plane and for assessment of cardiac pumping a short axis view with multiple slices covering the left ventricle is clinical standard with 3 standardized long axis views, called 2-chamber, 3-chamber and 4-chamber, as complement. **Figure 4** shows the three standardized long axis views and one slice of a short axis view of the left ventricle. Using cine bSSFP, CMR is the most accurate and reproducible method for assessing ventricular volumes and masses.

## Infarct imaging

Cardiovascular magnetic resonance is also gold standard for assessment of myocardial infarction using late gadolinium enhancement (LGE) which is based on using a gadolinium based contrast agent and a pulse sequence of inversion recovery(IR) [2-3, 34]. The gadolinium based contrast is injected intravenously and distributes in the extra cellular volume of all tissues and reaches a steady state after approximately 20 minutes. In an acute MI the cell membrane has ruptured and therefore the extra cellular volume is increased in comparison to normal myocardium and in a chronic MI the extra cellular volume is increased due to the scarred tissue. The gadolinium based contrast affect the T1 relaxation of the surrounding tissue to become shorter and thereby the regions of increased extracellular volume, as in the myocardial infarction, can appear brighter in T1-weighted images. By using an IR sequence the signal from tissues with a given T1-value can be nulled by choosing an appropriate inversion time (TI). The TI is chosen to null the signal of normal myocardium and the normal myocardium will hence appear black and the myocardial infarction will appear hyperenhanced, bright. Choosing the appropriate



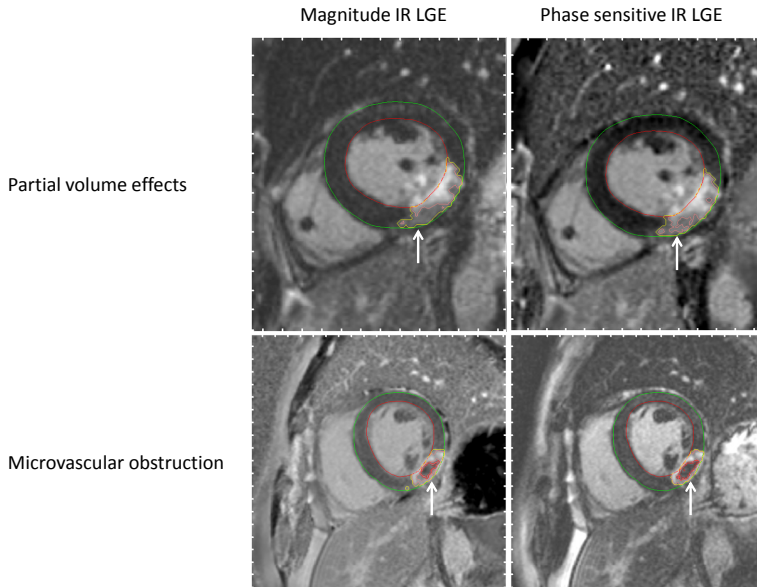
**Figure 4 Standardized view of the heart in cardiovascular magnetic resonance**

Three standardized long axis views of the human heart in cardiovascular magnetic resonance with a 2 chamber view (top left panel), 3 chamber view (top right panel) and a 4 chamber view (bottom left panel). The short axis view (bottom right panel) consists of multiple slices covering the whole heart.

inversion time is important to get a nulled signal from the myocardium and as good contrast as possible to the MI. The phase sensitive inversion recovery (PSIR) sequence was developed to decrease the sensitivity of TI and to increase image quality. Both IR and PSIR images have somewhat different characteristics. In usual magnitude IR images the nulled myocardium appears black and tissues with both shorter and longer T1 appears brighter. In PSIR images on the other hand the nulled myocardium is represented as gray and the hyperenhanced tissues with shorter T1 is brighter and the tissues with longer T1 are represented as darker than the normal myocardium. Both IR and PSIR LGE images are used as clinical standard and for visual assessment of myocardial infarction the intensity levels in the images are often set by the observer so that the normal myocardium is represented as black and the hyperenhancement represented as white. In these windowed images partial volume effects are then seen as a grey region between the bright and black region. Partial volume effects are due to the pixel volume containing more than one type of tissue for which the intensities are averaged and often seen at the border between different types of tissues. For myocardial infarction, however, there is a larger region seen as partial volume effects, representing partially infarcted regions. The partial volume effects representative of partially infarcted myocardium affects the quantification since the intensity does not represent fully infarcted regions of the myocardium. Also microvascular obstruction affects the quantification of myocardial infarction since it is seen as a darker region within the myocardial infarction and not shown as hyperenhancement despite being a region of no blood flow. In **Figure 5** examples of partial volume effects and microvascular obstruction in LGE images are shown.

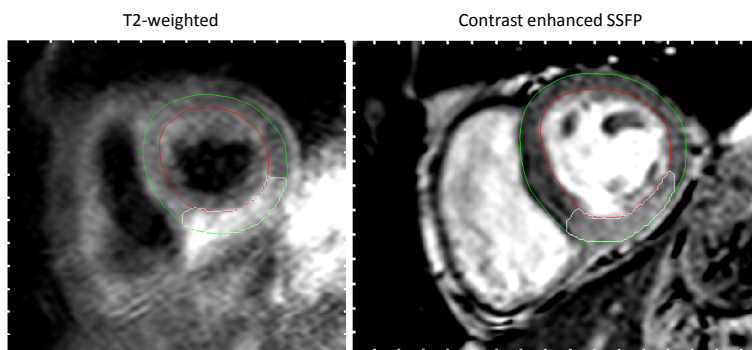
## Edema imaging

With CMR it is also possible to image edema [33] due to the increased water content in the tissue which increases the time for T2 relaxation. The extent of the edema in acute myocardial infarction has been shown to correlate well with the myocardium at risk defined by microspheres in dogs [33] and the pathologically defined myocardium at risk in pigs [35]. T2-weighted imaging has also been validated against SPECT in pigs as well as in patients and shown to depict the myocardium at risk (MaR) up to one week after acute MI [4]. Lately, also contrast enhanced SSFP (CE-SSFP) has been validated for assessment of myocardium at risk (MaR) up to one week after acute MI [5]. Five minutes after gadolinium injection a balanced SSFP imaging sequence is applied. In the edematous region the extracellular volume is increased and thereby the water content increases T2 and the gadolinium based contrast agent decreases the T1 relaxation. The SSFP sequence is affected by the ratio of T2/T1 and an increased signal can thus be seen in the edematous region. The two techniques of assessing MaR has been shown to correlate well in a direct comparison [36] and CE-SSFP has recently been shown to have a higher degree of diagnostic image quality across vendors and sites [21]. **Figure 6** shows an example of T2-weighted imaging and CE-SSFP in the same patient.



**Figure 5 Partial volume effects and microvascular obstruction in late gadolinium enhancement**

Myocardial infarction is not always seen as a compact region of increased signal intensity. Myocardial infarction can be affected by partial volume effects seen as a gray region, indicated by the arrows in the upper panels and by microvascular obstruction seen as a dark region within the myocardial infarction delineated in red and indicated by the arrows in the bottom panel. Myocardial borders are delineated in red (endocardium) and green (epicardium), myocardial infarction is delineated in yellow with a pink line indicating the core of the infarct and a red line indicating microvascular obstruction.



**Figure 6 Myocardium at risk as visualized in T2-weighted and contrast enhanced SSFP images**

Myocardium at risk can be visualized by cardiovascular magnetic resonance using either T2-weighted (left panel) or contrast enhanced SSFP images(right panel).Myocardium at risk is seen as a brighter region in the myocardium, and delineated in white. The myocardial borders are delineated in red (endocardium) and green (epicardium).

# Image segmentation

## Intensity classification

Image the intensity is the most important feature for automatic segmentation in CMR images. Analysis of the intensities is required both for simple threshold methods and as input to more advanced segmentation methods. Especially for segmentation of myocardium at risk and myocardial infarction simple threshold methods have been proposed to distinguish hyperenhanced myocardium from normal myocardium. The intensity thresholds that have been proposed are standard deviations from remote (n-SD), full width half maximum (FWHM) [22, 25], Otsu's threshold [17, 23] and Expectation Maximization (EM) [37-38]. The threshold methods of n-SD, FWHM and Otsu have been evaluated in several studies [18, 22, 27, 29] but there is no clinical standard [7] for quantification of myocardial infarction and myocardium at risk due to the differing results in the studies.

Using standard deviations from remote is a semi-automatic threshold method in which the user defines a remote region, representative of the normal myocardium. The threshold is then defined from the mean and standard deviation (SD) of the intensity of the remote region as  $\text{mean} + n\text{-SD}$ . The number of standard deviations,  $n$ , to use is dependent on the image quality regarding contrast intensity between normal and hyperenhanced myocardium, however this contrast may change between different MR scanners. Therefore, in addition to the difficulty in defining a remote region representative of normal myocardium, the n-SD threshold also has the drawback of not taking any information on the intensity of hyperenhancement into consideration when setting the threshold.

Amado et al. [22] proposed the threshold of FWHM in which the maximal intensity of hyperenhancement and the minimal intensity of the normal myocardium are used to define the threshold. The FWHM threshold is then set to half the distance between the minimum intensity of normal myocardium and maximal intensity of hyperenhancement. Thereby both the normal and hyperenhanced myocardium is taken into consideration. However, only minimal information on the intensity distributions are used with two intensity values and the FWHM threshold assumes equal standard deviation of the normal and hyperenhanced myocardium for the threshold to result in minimal erroneous classification. Hsu et al. [25] developed a more advanced segmentation algorithm, the FACT algorithm, for quantification of myocardial infarction in which a modified FWHM threshold was used to take partial volume effects of myocardial infarction into account. The modified FWHM threshold is defined from the mean intensity in a remote region and the maximum of the hyperenhancement.

Both the thresholds of n-SD from remote and FWHM from minimum intensity [22] and FWHM from mean remote intensity [25] needs user input and only a limited amount of information of the intensity distributions is utilized to define the threshold. Otsu's threshold [23] utilizes the whole intensity histogram to find a threshold. The threshold is found through exhaustive search of the histogram to find the threshold which minimizes the variance of the pixels below and above the threshold and thereby also maximizes the distance between the mean of the intensities above and below the threshold. As constructed Otsu's threshold will though favor equal distribution above and below the threshold and



this may result in close to 50 % myocardium at risk respectively myocardial infarction if there is a poor contrast between normal myocardium and hyperenhancement.

Expectation Maximization [37] is a statistical method which can be used to define an intensity threshold from the intensity histogram. Expectation Maximization was in this thesis implemented to assume a mixture of normal intensity distributions and estimate the mean and standard deviation of the intensity distributions. The maximum likelihood estimate of mean, standard deviation and the proportion of the distributions is iteratively estimated from an initial classification of the intensities as belonging to one or the other intensity distribution. To get an optimal threshold the initial estimation is of importance, especially if there is a small proportion for one of the intensity distributions for which there is a similar tendency to favor equal distribution above and below the threshold as for Otsu's threshold.

## Segmentation methods

More advanced methods for quantification of myocardial infarction [25-26] and myocardial edema [19-20] has been suggested by also incorporating post processing and regional analysis of the segmentations. For detection of microvascular obstruction both morphological operations [25] and a flood fill algorithm has been used [26]. Morphological operation is a method which can be specifically used to remove irregularities in the segmentation. Morphological operations shrink and expand objects and by combining these operations holes, as microvascular obstruction, and islands, as artifacts, can be removed from the segmentation. By using a flood fill method regions of the same intensity can be found and this can also be used to fill larger holes in the segmentation. The flood fill method is a segmentation method which finds connected regions of the same intensity and it can be used as a separate segmentation method or as post processing such as detection of microvascular obstruction.

Segmentation methods for the left ventricle segmentation methods are not only based on intensity thresholds and post processing but commonly based on more advanced segmentation methods [9]. Deformable models [39] and level set methods [40] are used to incorporate information on intensity and edges in the image and smoothness of the segmentation. This is achieved by modifying an initial segmentation to reach a position in which an artificial energy is minimized. The energy to minimize can be customized to the application at hand by which forces it is comprised of and their weights. The most important force is the mapping of the image intensities, often referred to as a balloon image in deformable models and a speed image in level set methods. The force based on intensities are then balanced by for example a curvature force, smoothing force and edge force. *A priori* information on the object to segment is most commonly incorporated in the deformable model or level set methods by user input or by creating a model of the object. Active shape models [41] is a specific type of deformable models for which a model of the object to segment is created from a training set which in the segmentation process is refined to fit the image. Active shape models incorporate a lot of *a priori* information, and can thus be used to design automatic algorithms for specific problems. However, active shape models require a training set of images with manual delineation of the object to segment and a drawback of the method is that it can be more difficult to perform segmentation of objects with a shape that was not represented in the training set.

## Validation

In the development of automatic algorithms for segmentation of the left ventricle a variety of different error measures have been used and the algorithms have been applied to different images. Manual delineations are used as reference for quantitative assessment of the automatic algorithms. The error to reference delineation as bias in clinical parameters as EDV, ESV, EF and LVM are reported for clinical relevance but also common image processing errors are commonly reported. In the review by Petitjean and Dacher [9] the perpendicular distance between the automatic and manual delineation, point-to-curve (P2C) was found to be the error most often reported for LV segmentation in CMR and was at the focus of their comparison. Petitjean and Dacher pointed out the difficulty in comparing the performance of the different segmentation methods due to the large variation in image quality and patient populations in which the algorithms have been applied. Two challenges, MICCAI'09<sup>1</sup> [9] and STACOM'11<sup>2</sup> [42], have been arranged for left ventricular segmentation in which patient data was supplied and a wide range of error measurements were computed. Amongst others the Dice similarity coefficient was used which was also suggested by Petitjean and Dacher [9] for future validation in combination with P2C and clinical parameters. By using DSC the overlap between the region of manual delineation and automatic segmentation is measured where, 1 indicate perfect overlap, 0 no overlap and DSC>0.7 is considered good [43]. Petitjean and Dacher [9] also concluded that the challenge databases should be used to overcome the differences in image quality. However, images in the MICCAI challenge were acquired without parallel imaging which is now clinical standard and in STACOM'11 not all results were derived using manual delineation as ground truth. Therefore a new training set and test set is needed for validation of left ventricular segmentation.

Validation of the more advanced algorithms for myocardial infarction has been performed both in animal studies [25-26] and in patient studies [26, 28] and simple threshold methods have been validated in animal studies [22] and patient studies with manual delineation as reference [18, 27]. Manual delineation has though been shown to have a large inter observer and intra observer variability [29] therefore validation in animal studies with an independent reference method is of importance [22, 25-26]. However, validation in patient studies is also necessary and this has only been performed in single center studies [26, 28] and thus it is not known if the results are applicable in other research and clinical centers.

For segmentation of myocardium at risk simple thresholds of 2-SD [14, 44], FWHM[17] and Otsu have been applied and compared to manual delineation [18, 45]. A more advanced method was developed by Johnstone et al. [19] who reported both bias to manual delineation and regional overlap by DSC. T2-weighted imaging has been validated in animal studies [33] but most importantly for validation of automatic segmentation methods both T2-weighted imaging and CE-SSFP have been validated against single photon emission computed tomography (SPECT), making it possible to perform independent validation in patient studies.

---

<sup>1</sup> <http://sourceforge.net/projects/cardiac-mr/files>

<sup>2</sup> <http://www.cardiacatlas.org>

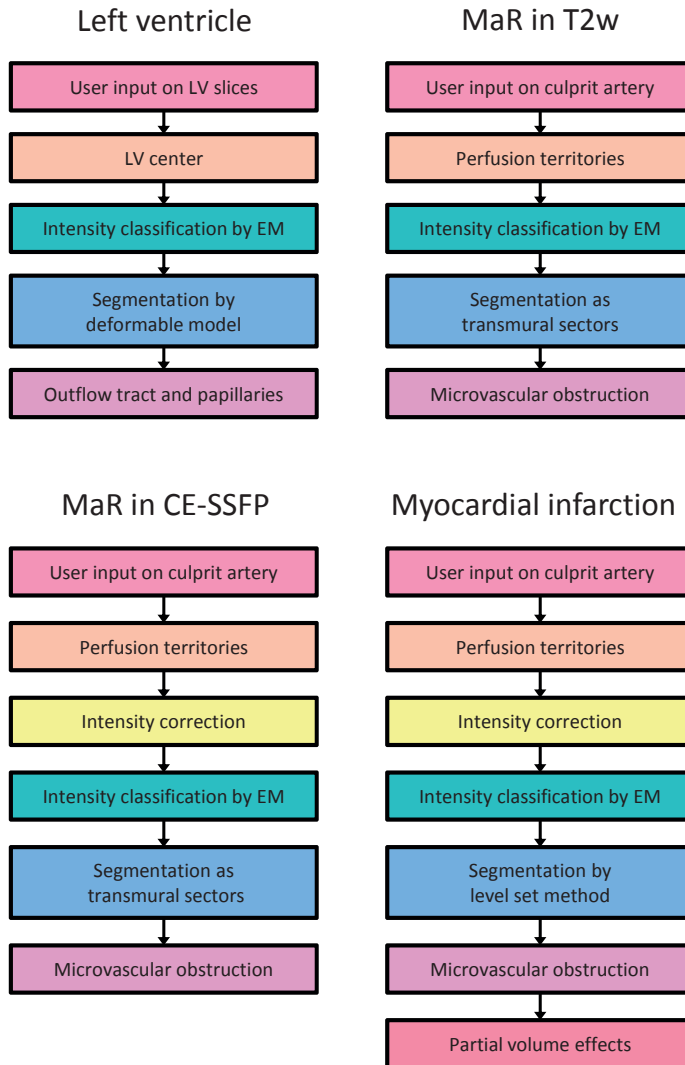


# Methods

The new automatic algorithms were all implemented using *a priori* information for pre and post processing in combination with intensity classification by Expectation Maximization (EM). All algorithms were validated against reference delineation in patient data and where available to an independent reference standard. Further, the proposed algorithms were compared to previously suggested segmentation methods.

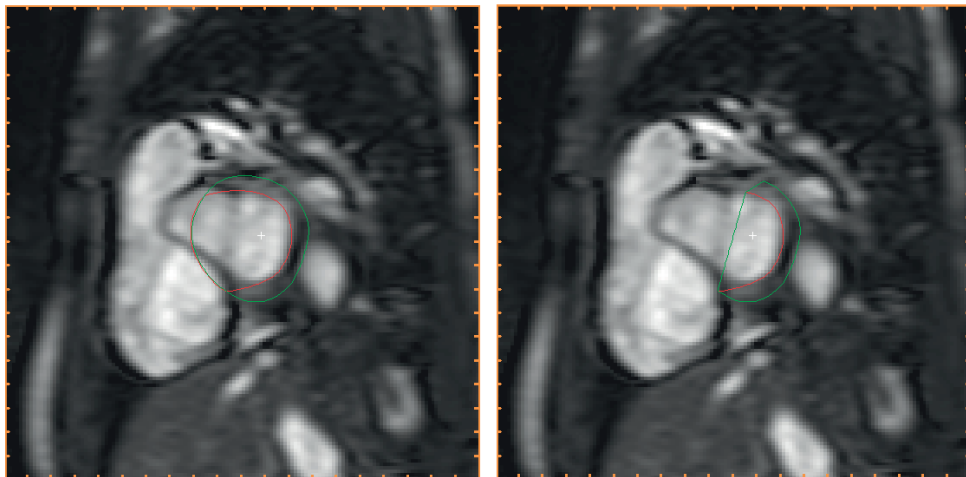
## Automatic segmentation algorithms

Four automatic algorithms were developed, segmentation of the left ventricle in cine SSFP images (Paper II), quantification of myocardium at risk in T2-weighted images (Paper III) myocardium at risk in contrast enhanced SSFP images (Paper IV), and finally quantification of myocardial infarction in LGE images (Paper V). All four algorithms are based on using Expectation Maximization (EM) for intensity classification and utilize *a priori* information for pre and post processing. Simplified flow charts of the algorithms are shown in **Figure 7**. All four algorithms require initial user input, for left ventricular segmentation (Paper II) slices containing any myocardium are selected after which the LV center point is automatically detected, for segmentation of myocardium at risk and myocardial infarction (Paper III-V) the user indicates culprit artery and rotation of the LV after which *a priori* information on perfusion territories of the coronary arteries can be applied. Surface coil intensity correction was implemented for segmentation of myocardium at risk in CE-SSFP (Paper IV) and segmentation of myocardial infarction in LGE images (Paper V). Intensity classification by EM was used with initializations and constraints specifically designed for each of the segmentation algorithms. Segmentation was performed using a deformable model for the left ventricle, using a model of transmural, connected myocardial sectors for myocardium at risk and using a level set method for infarct quantification. Post processing was performed using *a priori* information on left ventricular function to detect outflow tract and papillaries (Paper II) and *a priori* information on microvascular obstruction (Paper III-V). In Paper V partial volume effects were taken into consideration by using a weighted summation of pixel intensities. All four algorithms were implemented in the software Segment (<http://segment.heiberg.se>) which is described in Paper I and freely available for research.



**Figure 7 Flow chart of the four algorithms implemented in Paper II-V**

Flow chart of the four automatic algorithms, segmentation of the left ventricle (LV) in SSFP images (Paper II), myocardium at risk (MaR) in T2-weighted images (Paper III) and contrast enhanced SSFP (CE-SSFP) images (Paper IV) and myocardial infarction in late gadolinium enhancement (LGE, Paper V). All algorithms are shown with simplified processing blocks. Intensity classification by Expectation Maximization (EM, green) is the basis for all four algorithms with different segmentation methods (blue) and *a priori* information utilized in both pre processing (orange), intensity correction (yellow) and post processing (purple). User input is shown in pink and the weighted summation of pixel intensities to account for partial volume effects in segmentation of myocardial infarction is shown in red.



**Figure 8** Detection of outflow tract in left ventricular segmentation

The left ventricular segmentation needs to be adjusted in the basal slices using *a priori* information on the outflow tract moving in and out of the imaging plane during a heartbeat. Endocardial (red) and epicardial (green) segmentations are shown in one basal slice in end diastole prior to detection of the outflow tract (left panel) and after adjustment for the outflow tract (right panel).

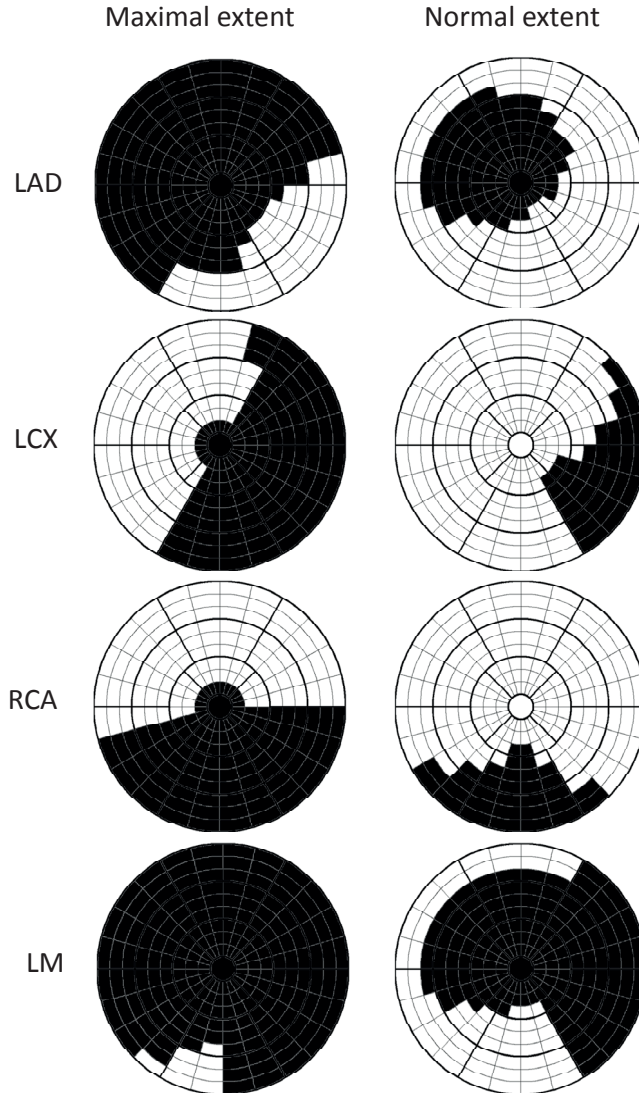
## *A priori* information

### *Left ventricular function*

For segmentation of the left ventricle it is important to get endocardial and epicardial segmentation as close as possible to reference delineation and to detect both papillary muscles and outflow tract. *A priori* information was incorporated by using a training set of 40 subjects for optimization of parameters in the segmentation algorithm. For inclusion of papillary muscles in the endocardial segmentation a convex hull was used and the endocardial segmentation was further expanded by assuming constant papillary volume over time. In basal slices the outflow tract was detected as sectors for which the intensities were not typical of myocardium and as sectors with a mean wall thickness less than 2 millimeters. Intensities were considered not typical for myocardium if the intensities were above mean + two standard deviations of the myocardial intensities. **Figure 8** shows an endocardial and epicardial segmentation pre and post detection of the outflow tract.

### *Coronary artery perfusion territories*

For segmentation of myocardium at risk and myocardial infarction an extent model of the perfusion territories of the coronary artery was used as *a priori* information and the user selected culprit artery as input to the automatic algorithm. Normal and maximal extent for the perfusion territories of each coronary artery was defined in consensus by three experienced observers from their combined experience of CMR and SPECT. Normal and maximal extent models were defined by each observer and discussed until consensus was reached for left anterior descending artery (LAD), left circumflex artery (LCx) and right



**Figure 9** *A priori* information on perfusions territories

Bulls-eye representation of maximal extent model (left column) and normal extent model (right column) for the perfusion territories of left anterior descending artery (LAD), left circumflex artery (LCx), right coronary artery (RCA), and left main artery (LM). Models for LAD, LCX and RCA were defined in consensus by three experienced observers in an extended 17-segment AHA model and models for LM were defined from the models of LAD, LCX and RCA. The 17-segment model is extended to three slices in each of the basal, mid-ventricular and apical zones and 24 sectors in each slice. Black sectors are included in the maximal and normal extent model, respectively. The septal part of the left ventricle is represented in the left of the bulls-eye plot, the lateral part in the right, anterior part in the top, inferior part in the bottom, the apical slices in the center and the basal slices in the outer part of the bulls-eye plot.

coronary artery (RCA). The models for left main artery (LM) were defined from the models of LAD, LCx and RCA. The extent was defined in the 17 segment American Heart Association (AHA) model [46] which was extended to 24 sectors circumferentially and 3 slices in each of the basal, midventricular and apical parts of the left ventricle. The maximal and normal extent models as defined in consensus are shown in **Figure 9**.

The normal extent model was only used for initialization of the EM-algorithm in segmentation of myocardium at risk in T2-weighted images (Paper III). The maximal extent model was used for initialization of the EM-algorithm in segmentation of myocardium at risk in CE-SSFP (Paper IV) and to remove artifacts outside the maximal extent of the perfusion territory (Paper III-V). The maximal extent model was further used to define remote myocardium for surface coil intensity correction in CE-SSFP and LGE images (Paper IV and V).

### *Microvascular obstruction*

Microvascular obstruction is in LGE images as well as in T2-weighted and CE-SSFP images seen as a hypo enhanced region within the myocardial infarction. Microvascular obstruction should be included in the region of myocardial infarction respectively myocardium at risk despite being manifested as a dark region [32]. For myocardium at risk (Paper III and IV) segmentation was performed sector wise and myocardium at risk was required to be transmural. Microvascular obstruction was detected as missing sectors within the region of connected sectors by use of morphological closing. For segmentation of myocardial infarction (Paper V) transmural is not required and microvascular obstruction was instead detected as a hole within the region of infarct by using a flood fill algorithm. The flood fill algorithm was also combined with morphological closing to detect large regions of microvascular obstruction with a thin border of surrounding hyper enhanced myocardial infarction.

### *Surface coil intensity correction*

Varying surface coil sensitivity causes an intensity gradient through the CMR images. This can cause a larger variability in the myocardium than the contrast between normal myocardium and hyper enhancement representative of myocardium at risk or myocardial infarction. In CE-SSFP and LGE images (Paper IV and V) a surface coil intensity correction was implemented based on the intensity of the blood pool and the remote myocardium, defined by the maximal extent model. A second order linear intensity correction was applied to account for a gradient proportional to the squared distance to the surface coil. The intensity correction should result in a reduced variability in the remote region and a mean intensity in the culprit region higher than in the remote region, otherwise the correction was not applied. Intensity correction was not applied in T2-weighted images (Paper III) due to the black blood property and instead intensity classification was applied slice by slice to overcome the intensity gradient across slices.

### *Partial volume effects*

Partial volume effects can be seen in all CMR images due to limited spatial resolution, especially in through plane direction but can also be due to partially infarcted regions as in LGE images (Paper V). To account for partial volume effects in LGE images Heiberg et al. [26] calculated scar size as a weighted sum with weights representative of the amount of scarred cells within the pixel. Heiberg et al. [26] though used a fixed threshold of 1.8SD to

define the infarct region and a fixed minimum weight of 0.1 at the minimum intensity. In Paper V an adaptable threshold was implemented by EM and hence adaptable weights were also needed. The weights for normal myocardium should be set to 0 for pixels representative of normal myocardium and 1 for pixels representative of fully scarred myocardium. Fully scarred myocardium was chosen as the 90th percentile of the pixels within the scar region and weights were calculated as a linear function from 0 at the mean intensity of normal myocardium.

## Intensity classification by Expectation Maximization

For classification of pixel intensities a Bayesian probability is calculated by use of an Expectation Maximization algorithm (EM) [37]. The EM-algorithm iteratively refines an initial classification to find the maximum likelihood estimate of mean and standard deviation for the assumed Gaussian intensity distributions. For endocardial segmentation (Paper II) EM was used to estimate the intensity distributions of blood and myocardium and for epicardial segmentation the intensity distribution of surrounding tissues was also estimated. For segmentation of myocardium at risk (Paper III and IV) the intensity distributions of normal myocardium and myocardium at risk were estimated and for myocardial infarction (Paper V) the intensity distributions of normal myocardium and myocardial infarction were estimated. Intensity classification by EM can be sensitive to the initial classification and it was therefore chosen carefully. In LV segmentation (Paper II) the initial classification between blood and myocardium was set as two standard deviations above the mean intensity of a cylinder with radius 10 mm placed at the LV center point in the midventricular slices. For segmentation of myocardium at risk in T2-weighted images (Paper III) the normal extent model of the culprit artery was used to initialize all pixels within the normal extent as myocardium at risk and all pixels outside as normal myocardium. For segmentation of myocardium at risk in CE-SSFP (Paper IV) the initial classification was modified to use the maximal extent model instead to get an initial classification of normal myocardium as representative as possible. In paper IV a larger overlap between normal myocardium and myocardium at risk was seen and a constraint of keeping initial classifications was presented. Pixels with intensity below the 50th percentile in the remote region were kept classified as normal myocardium and intensities above the 75th percentile in the culprit region were kept classified as myocardium at risk. For infarct segmentation (Paper V) initialization of EM was based on assuming a small myocardial infarction represented as the intensities above the 90th percentile of the intensity histogram. The constraints was set to keep the intensities below the 5th percentile of the remote region classified as normal myocardium and pixels above the 95th percentile were kept classified as myocardial infarction.

# Validation

## Software for cardiac image analysis (Paper I)

Validation of the Segment software was performed by implementing an extensive test script that runs on archived test data. The test script output is then compared with known accurate results from previously validated scientific publications.

## Left ventricular segmentation (Paper II)

In total 90 subjects were included in the study, both patients referred for clinical evaluation of known or suspected coronary artery disease as well as healthy subjects and athletes. The subjects were scanned using a 1.5T MR scanner (Philips Intera CV, Philips, Best, The Netherlands). The sequence used was a balanced steady state free precession (bSSFP) sequence with retrospective ECG triggering.

The subjects were divided into a training set (n=40) and a test set (n=50). The test set consisted of 20 patients, 20 healthy volunteers and 10 athletes. Manual segmentation was performed for all slices in end-diastole and end-systole by an experienced observer. A subset of 25 subjects from the test set (10 patients, 10 healthy volunteers and 5 athletes) were used for second observer analysis of manual delineation.

In the test set the difference between manual delineation and automatic segmentation was computed for the clinical parameters end-diastolic volume (EDV), end-systolic volume (ESV), ejection fraction (EF) and left ventricular mass (LVM) as well as the image processing error measurements of dice similarity coefficient (DSC) [47] and point to curve distance (P2C). All errors were reported as mean  $\pm$  standard deviation. Calculation of both clinical parameters and image processing errors allowed for comparison against previous methods and both training set and test set are available upon request for benchmarking of future algorithms.

## Myocardium at risk in T2-weighted CMR (Paper III)

Forty seven patients with first time ST-elevation myocardial infarction (STEMI) treated with primary percutaneous coronary intervention (PCI) were included in this study. Within a week after admission patients were imaged using 1.5T scanner from Siemens (Magentom Vision, Siemens, Erlangen, Germany) or Philips (Philips Intera CV or Achieva, Philips, Best, The Netherlands). The sequence used was a T2-weighted triple inversion turbo spin echo sequence (T2-STIR). Regions of hyper enhanced myocardium was manually delineated as myocardium at risk (MaR) by an experienced observer and expressed as percent of left ventricular mass (%LVM). Hypo-intense regions within MaR was considered microvascular occlusion and included as MaR.

Quantification of MaR by the new automatic algorithm, manual second observer delineation and the threshold methods of 2SD from remote, FWHM from remote as



implemented by Hsu [25] and Otsu's threshold [23] were all compared against the reference observer using Bland-Altman bias (mean  $\pm$  standard deviation) and linear regression analysis (correlation coefficient). Regional agreement to reference observer was evaluated as DSC (mean  $\pm$  standard deviation) [47].

## Myocardium at risk in CE-SSFP (Paper IV)

For validation of the new automatic algorithm 183 patients from the clinical cardio protection trials CHILL-MI [48] (n=92) and MITOCARE [49] (n=91) were included. Patients who had undergone CMR examination 2-6 days following first time STEMI treated with PCI with LGE images and CE-SSFP images of diagnostic quality were included. For validation against an independent reference method, an additional set of patients who had undergone both CE-SSFP CMR and single photon emission computed tomography (SPECT) [5] (n=16) were included in this study.

All CMR examinations were performed on 1.5T scanners from Philips (Philips healthcare, Best, The Netherlands), Siemens (Siemens AG, Erlangen, Germany) or GE (GE Healthcare, Waukesha, WI, USA). Cine CE-SSFP images were obtained approximately 5 minutes after injection of an extracellular gadolinium -based contrast agent and LGE images were acquired approximately 15 minutes after injection of the gadolinium-based contrast agent. Surface coil intensity correction was not generally applied across vendors and sites. Myocardial perfusion SPECT was performed in the additional set of 16 patients within four hours after PCI using either of two dual head cameras GE (Ventri, GE Healthcare, Waukesha, WI, USA) or Sopha (DST-XL, Sopha Medical Vision, Bue, Cedex, France).

In CE-SSFP images, hyper-intense regions within the myocardium were manually delineated as MaR and in LGE images, hyper-intense regions were delineated as myocardial infarction using a previously validated automatic segmentation algorithm [26], followed by manual corrections when needed. Hypo-intense myocardium within the area of increased signal intensity was regarded as microvascular obstruction and thus included. The delineation of each data set was performed by one of three experienced observers with a quality control of the delineations by a second opinion for each case. Different opinions for the delineation were resolved in consensus between all three observers when necessary. In a subset of 15 patients from the multi-center studies, second observer manual delineation of MaR was performed to evaluate inter-observer variability. In SPECT images, MaR was delineated by use of an 55 % threshold [50] and manual corrections after automatic delineation of epicardial and endocardial borders [51]. MaR was expressed as percent of left ventricular mass (%LVM).

Bland-Altman bias (mean  $\pm$  standard deviation), linear regression analysis (correlation coefficient) and regional agreement DSC to manual delineation was evaluated for the new automatic algorithm, and the threshold methods of 2SD from remote, FWHM from remote [25] and Otsu's threshold [23] as well as for inter-observer analysis. Quantification of MaR in CE-SSFP images by the new automatic algorithm and manual delineation was compared to quantification of MaR in SPECT using bias (mean  $\pm$  standard deviation) and linear regression analysis (correlation coefficient). The added value of the processing blocks in the new automatic algorithm was assessed by both bias and DSC and evaluated using two-sided paired t-test with Bonferroni correction.



## Myocardial infarction (Paper V)

The new automatic algorithm was validated experimentally against TTC and *ex-vivo* high resolution T1-weighted images. Pigs with induced myocardial infarction was included from three previous studies, one mechanistic basic research study [52], one cardio protection study [53] and one cardio protection study also used for validating the original weighted algorithm for infarct quantification [26]. All pigs were subjected to 40 minutes occlusion with a balloon placed after the first or second diagonal branch of the LAD. Myocardial infarction was imaged after four hours [26, 53], six hours [52] or seven days of reperfusion [52] with one or more of the following CMR images, *in-vivo* 3D IR LGE (n=20), *in-vivo* 2D PSIR LGE (n=12) and *ex-vivo* high resolution T1-weighted images (n=38). Six pigs with myocardial infarction were imaged after seven days of reperfusion with all three CMR images and following *ex-vivo* imaging, hearts were sliced into five mm slices and incubated in triphenyltetrazolium-chloride (TTC) for five minutes. CMR imaging was performed on a 1.5T Philips scanner (Philips Healthcare, Best, The Netherlands). *In-vivo* LGE images were acquired approximately 20 minutes after injection of gadolinium-based contrast agent. *Ex-vivo* high resolution T1-weighted images were acquired with the explanted hearts placed in plastic containers and the ventricles filled with balloons containing deuterated water.

The new automatic algorithm was also applied in multi-center, multi-vendor patient studies with expert delineation as reference. Patients with first time ST-elevation myocardial infarction (STEMI) treated with percutaneous coronary intervention (PCI) were included from the recently published clinical cardioprotection trials CHILL-MI [48] and MITOCARE [49]. All CMR examinations were performed on 1.5T scanners from Philips (Philips Healthcare, Best, The Netherlands), Siemens (Siemens AG, Erlangen, Germany) or GE (GE Healthcare, Waukesha, WI, USA). LGE images covering the entire left ventricle were acquired approximately 15 minutes after injection of the gadolinium-based contrast agent using inversion recovery gradient-recalled echo sequence. Inversion time was manually adjusted to null the signal of viable myocardium and surface coil intensity correction was not generally applied across vendors and sites. Patients with LGE images of diagnostic quality acquired with IR (n=75) or paired IR and PSIR images (n=49) were included in this study.

In LGE images, reference delineation of myocardial infarction was performed using the weighted method based on 1.8SD [26], with manual corrections by an experienced observer when needed. Hypo-intense regions within the region of gadolinium enhancement were considered as micro vascular obstruction and hence included in the analysis as 100 % scar. In the patient data, delineation of each data set was performed by one of three experienced observers with a quality control of the delineations by a second opinion and different opinions resolved in consensus between all three observers. In a subset of 17 patients a second observer delineation of myocardial infarction was performed to evaluate inter-observer variability of the reference delineation. In experimental LGE images delineation was performed by one experienced observer using the same methodology in LGE images [26] and using a threshold of 8SD from remote in T1-weighted images [26], with manual corrections where needed and hypo intense regions included as microvascular obstruction. In TTC images myocardial scar was manually delineated as the non-stained parts of the myocardium.

Validation was performed for *in-vivo* IR, *in-vivo* PSIR, and *ex-vivo* high resolution T1-weighted images against TTC for myocardial infarction imaged after seven days of reperfusion and against *ex-vivo* high resolution T1-weighted images. The new automatic algorithm was applied in multi-center, multi-vendor patient data with expert delineation as reference. Comparisons were performed using bias (mean  $\pm$  standard deviation) and linear regression analysis (correlation coefficient). The applicability of the new automatic algorithm was assessed in paired IR and PSIR images using bias (mean  $\pm$  standard deviation), linear regression analysis (correlation coefficient) and regional agreement DSC (mean  $\pm$  standard deviation) for the new automatic algorithm, the original weighted algorithm and the threshold methods by EM, 2, 3 and 5 SD from remote, FWHM from minimum intensity as implemented by Amado et al. [22], FWHM from mean remote intensity as implemented by Hsu et al. [25], and Otsu's threshold [23].

# Results and comments

## Software for cardiac image analysis (Paper I)

The Segment software is a comprehensive software for cardiovascular image analysis which has been used in a wide range of publications. The software includes loading of DICOM images, image display, flow quantification with automatic vessel segmentation, automatic segmentation of the left ventricle and myocardial infarction, tools for manual delineation and a tool for general object segmentation. The Segment software was extended with possibility to write own dedicated plug-ins in Matlab and an automatic test script for validation. The test script is continuously refined as new features are added to the software and the output of the test script is compared to known accurate results from previously validated scientific publications.

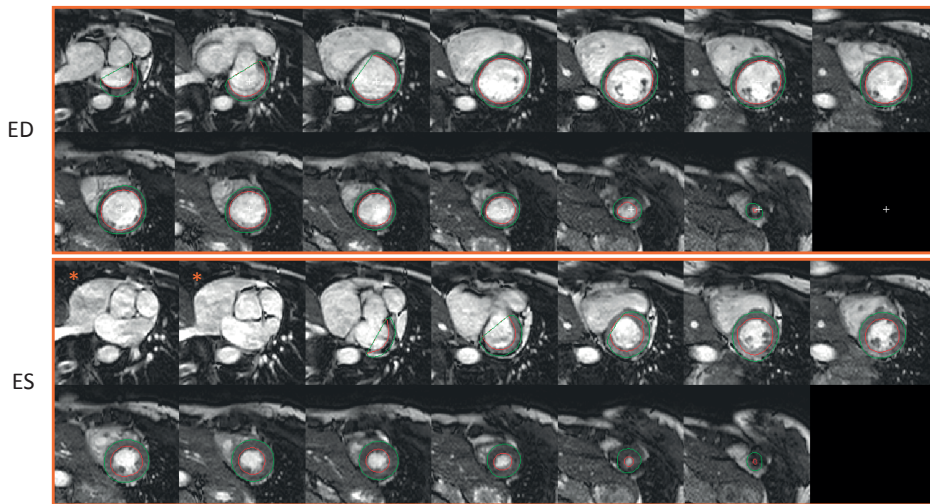
The automatic test script enabled an FDA 510 (k) application which was approved and the software is being used in hospitals both in Europe and USA. The test script is continuously expanded as the Segment software is improved and new algorithms developed. The algorithms developed in this thesis are all implemented in the software and made freely available upon publication of the papers.

## Left ventricular segmentation (Paper II)

The new automatic left ventricular segmentation was validated against manual delineation in a test set of 49 subjects and compared to inter observer variability in a subset of 24 patients. A typical example is shown in end diastole and end systole in **Figure 10**. Bias to manual delineation in the test set was EDV  $-11 \pm 11$  ml ( $R=0.96$ ), ESV  $1 \pm 10$  ml ( $R=0.95$ ), EF  $-3 \pm 4$  % ( $R=0.86$ ) and LVM  $4 \pm 15$  g ( $R=0.87$ ) (**Figure 11**). Regional agreement to manual delineation was for endocardial segmentation DSC  $0.91 \pm 0.03$  and for epicardial segmentation DSC  $0.93 \pm 0.02$ . Inter observer variability of manual delineation in the second observer subset was EDV  $10 \pm 4$  ml, ESV  $5 \pm 5$  ml, EF  $0 \pm 2$  % and LVM  $-7 \pm 9$  g.

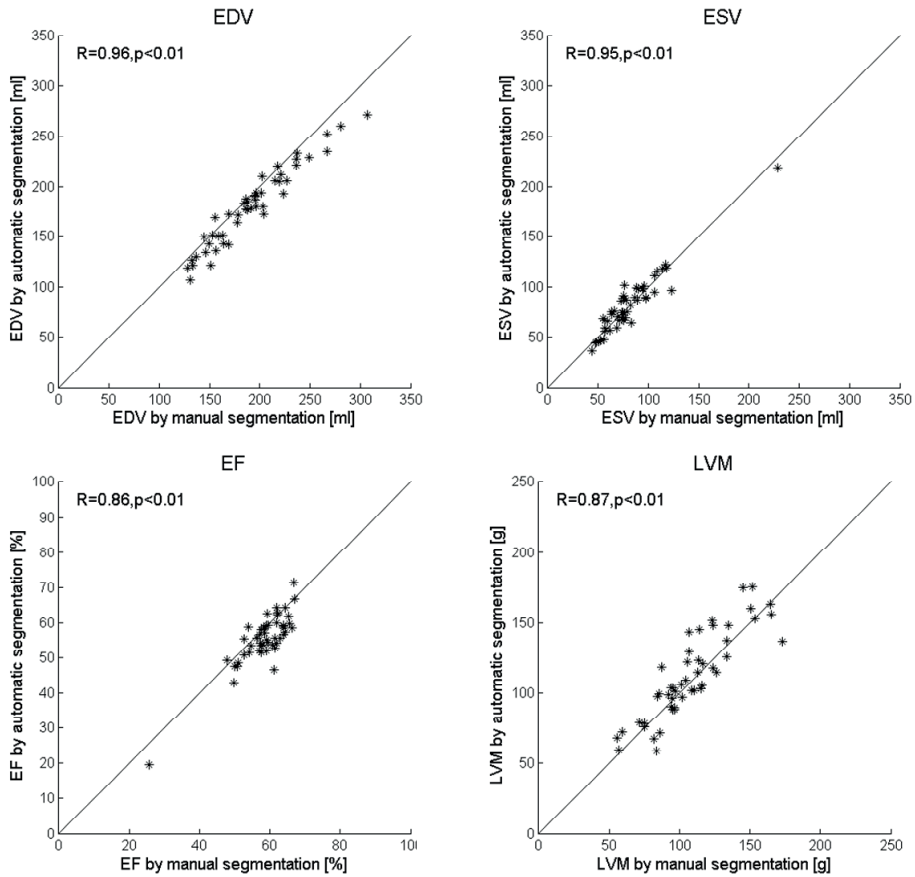
These results can be compared to the recent study by Suinesiaputra et al. [6] with manual delineation by seven expert observers from different sites compared to an algorithmically created consensus for 15 patients. A large range of bias to consensus was found for the seven observers EDV  $-37$  to  $40$  ml, ESV  $-33$  to  $41$  ml, EF  $-12$  to  $13$  % and LVM  $-44$  to  $60$  g and the average precision amongst the observers was EDV  $9$  ml, ESV  $9$  ml, EF  $4$  % and LVM  $11$  g, similar to the variability of the new automatic algorithm.

In paper I the Segment software was presented with an automatic algorithm for segmentation of the left ventricle implemented by Heiberg et al. [24] in 2005 and bias to manual delineation was  $EDV -1 \pm 11\text{ml}$  ( $R^2=0.99$ ) and  $LVM 4 \pm 15\text{ g}$  ( $R^2=0.94$ ). Variability of EDV and LVM for the new automatic algorithm was similar to the previous algorithm whereas bias was higher for the new automatic algorithm, however, ESV and EF was not reported for the previous algorithm and the new automatic algorithm included detection of the outflow tract (**Figure 10**) and was validated in a larger set of data (**Figure 11**).



**Figure 10** Example of segmentation in end-diastole and end-systole.

An example of automatic segmentation is shown in end-diastole (top panel) and end-systole (bottom panel). Each panel shows the short axis stack covering the left ventricle from base to apex with endocardial (red) and epicardial (green) segmentations. Note how the outflow tract has moved out of the two most basal slices in end-systole (bottom panel, images marked \*), compared to end-diastole (top panel) and that the algorithm has automatically corrected for this long-axis motion



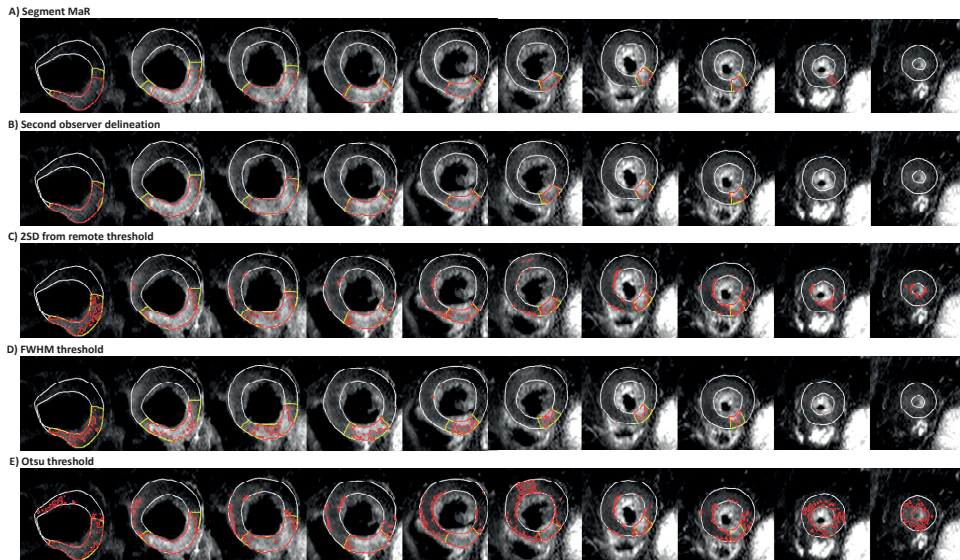
**Figure 11 Validation of automatic left ventricular segmentation**

Agreement between automatic segmentation and manual delineation of the left ventricle. Automatic segmentation plotted against end-diastolic volume (EDV, top left), end-systolic volume (ESV, top right), ejection fraction (EF, bottom left) and left ventricular mass (LVM, bottom right) for 49 patients in the test set. Solid lines indicate line of identity.

## Myocardium at risk (Paper III and IV)

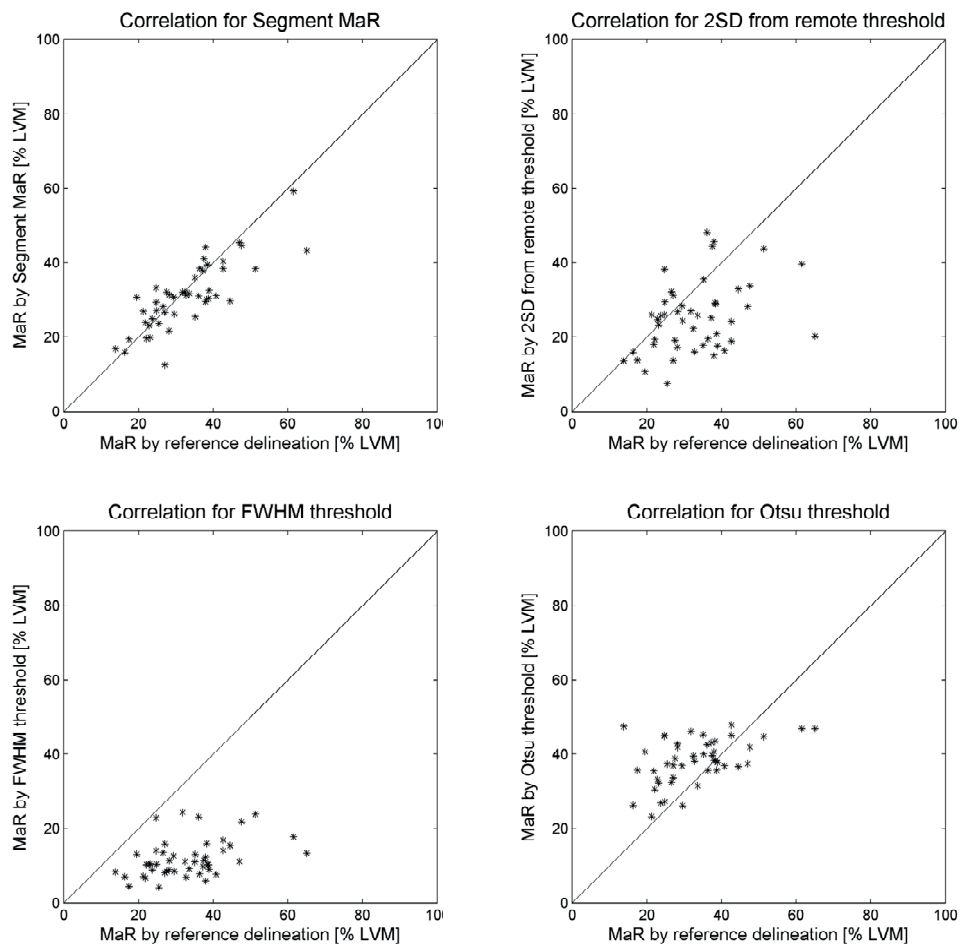
In paper III and IV automatic segmentation of myocardium at risk was developed for T2-weighted images and CE-SSFP images and compared to the previously suggested threshold methods. **Figure 12** shows a typical example of automatic segmentation of MaR in T2-weighted images in comparison to inter observer variability of manual delineation and the threshold methods of 2SD from remote, FWHM from mean remote intensity [25] and Otsu's threshold [23].

In paper III the automatic algorithm was validated against manual delineation in 47 patients with T2-weighted images from a single center with two different scanner vendors with a bias to manual delineation of  $-2 \pm 6\%$  LVM ( $R=0.81$ ) and regional agreement of DSC  $0.85 \pm 0.07$  (**Figure 13**). Inter observer analysis for manual delineation was performed in all 47 patients with a bias of  $-2 \pm 5\%$  LVM with regional agreement DSC  $0.90 \pm 0.08$ . In paper IV the automatic algorithm was validated in 183 patients with CE-SSFP images from multi-center, multi-vendor studies with a bias to manual delineation of  $1 \pm 6\%$  LVM ( $R=0.83$ ) and regional agreement DSC  $0.85 \pm 0.08$  (**Figure 14**). Inter observer analysis of manual delineation was performed in 15 patients with a bias of  $0 \pm 3\%$  LVM and regional agreement DSC  $0.92 \pm 0.04$ . Bias against SPECT was  $2 \pm 7\%$  LVM ( $R=0.73$ ) for the new automatic algorithm in comparison to  $1 \pm 5\%$  LVM ( $R=0.90$ ) for the manual delineation in CE-SSFP.



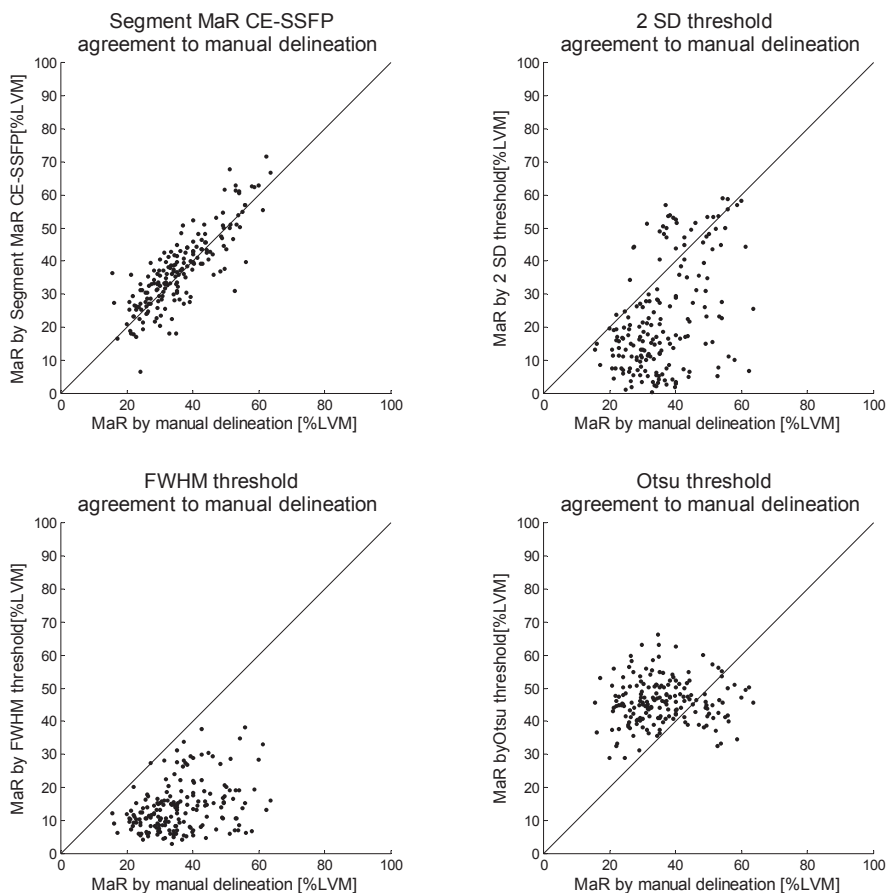
**Figure 12** Typical segmentation result of myocardium at risk in T2-weighted images

Typical MaR segmentation shown in red for the automatic Segment MaR (panel A), second observer delineation (panel B), the threshold methods of two standard deviations from remote (2SD, panel C), full width half maximum (FWHM, panel D) and Otsu (panel E), compared to manual delineation by the reference observer, shown in yellow. The most basal slice is shown in the left of the panel and the most apical slice in the right of the panel. Note the continuous appearance of the segmentation for Segment MaR and manual delineation compared to the threshold methods of 2SD, FWHM and Otsu.



**Figure 13 Validation of myocardium at risk in T2-weighted images.**

Agreement to manual delineation of myocardium at risk (MaR) as % of left ventricular mass (%LVM) for the automatic algorithm Segment MaR (top left), and the threshold methods of two standard deviations from remote (2SD, top right), full width half maximum(FWHM, bottom left) and Otsu's threshold (bottom right). Solid lines indicate line of identity.



**Figure 14 Validation of myocardium at risk in CE-SSFP images.**

Agreement to manual delineation of myocardium at risk (MaR) as % of left ventricular mass (%LVM) for the automatic algorithm Segment MaR CE-SSFP (top left), and the threshold methods of two standard deviations from remote (2SD, top right), full width half maximum (FWHM, bottom left) and Otsu's threshold (bottom right). Solid lines indicate line of identity.



In both Paper III and Paper IV the threshold methods of 2SD, FWHM from remote and Otsu's threshold showed a higher bias, lower regression R-value and poorer regional agreement DSC than for the respective automatic algorithms of Segment MaR T2 and Segment MaR CE-SSFP (**Table 1**). The performance of both automatic algorithms was evaluated step by step or block by block. In paper III all steps of the algorithm was shown to decrease bias except the final step which smoothed the MaR region across slices to get a physiological appearance. In paper IV not all processing blocks of the algorithm decreased bias to manual delineation however regional agreement DSC was increased in each block. The results in Paper IV thereby highlighted the importance of using both regional agreement and bias of clinical parameters to evaluate performance. **Figure 15** shows the block by block analysis of the new automatic algorithm in Paper IV.

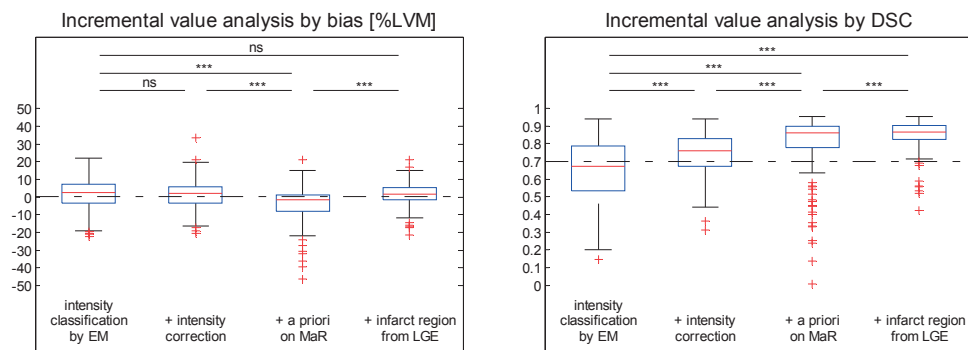
The first block of both methods was to use intensity classification by EM. However, it was only in Paper IV that the pure threshold by EM was superior to the thresholds of 2SD, FWHM and Otsu. This is probably due to the different constraint on the EM classification for T2-weighted images and CE-SSFP images. The constraint for T2-weighted images was set to keep initial classification of 80% of the pixels within the normal extent model as myocardium at risk whereas for segmentation of MaR in CE-SSFP the top 25% of the pixels within the maximal extent model were kept classified as MaR. For segmentation in T2-weighted images the MaR region is thereby likely to initially be overestimated. However, the remainder of the algorithm was built upon this threshold and a low bias and high regional agreement was found against manual delineation. The approach of initially overestimating a region and then applying *a priori* information has also been successfully implemented by Hsu et al. [25] for infarct segmentation in LGE images.

Both algorithms for segmentation of myocardium at risk (Paper III-IV) showed a low bias and high regional agreement to manual delineation; close to inter observer variability for manual delineation by expert observers. A limitation to Paper III was that T2-weighted images from a single center were used and no independent reference method was used whereas, in Paper IV multi-center, multi-vendor CE-SSFP images were used and the automatic algorithm was evaluated against SPECT with equally low bias and high regional agreement.

	T2-weighted			CE-SSFP		
	Bias to manual delineation [%LVM]	Regression R-value	DSC	Bias to manual delineation [%LVM]	Regression R-value	DSC
<b>Segment MaR</b>	-2 ± 6	0.81	0.85 ± 0.07	1 ± 6	0.83	0.85 ± 0.08
<b>2SD threshold</b>	-8 ± 11	0.38	0.69 ± 0.14	-13 ± 15	0.47	0.54 ± 0.27
<b>FWHM threshold</b>	-21 ± 10	0.41	0.46 ± 0.14	-22 ± 11	0.42	0.42 ± 0.21
<b>Otsu threshold</b>	5 ± 10	0.47	0.68 ± 0.10	10 ± 12	0.05	0.65 ± 0.12

**Table 1 Validation of segmentation of myocardium at risk in T2-weighted and CE-SSFP images**

Bias as % of left ventricular mass (%LVM), regression R-value and regional agreement as Dice similarity Coefficient (DSC) for segmentation of myocardium at risk in T2-weighted (n=47 patients) and CE-SSFP images (n=183 patients) for Segment MaR, and thresholds by two standard deviations from remote (2SD), full width half maximum (FWHM) and Otsu's threshold. Results are expressed as mean ± standard deviation.



**Figure 15 Block by block analysis with both bias and DSC**

Block by block analysis of the new automatic algorithm for MaR quantification in CE-SSFP images. Incremental value of each block in the automatic segmentation algorithm analyzed by bias to manual delineation as %LVM (left panel) and by regional agreement as Dice similarity coefficient DSC (right panel). Bias and DSC was calculated with segmentation based on only intensity classification by Expectation Maximization and calculated after the addition of the processing blocks of intensity correction, *a priori* on myocardium at risk (MaR) and infarct region from late gadolinium enhancement (LGE). For each block of the algorithm the upper limit of the box indicate upper quartile, middle line indicate median, lower limit of box indicate lower quartile, whiskers indicate minimum and maximum and points (+) indicate outliers. Bias zero is shown as dotted black line in the left panel, DSC above of 0.7 indicates good regional agreement, and is shown as dotted black line in the right panel. Two sided paired t-test was performed for each block in comparison to previous block and first block, ns: non significant, \*\*\*: p<0.0001.

## Myocardial infarction (Paper V)

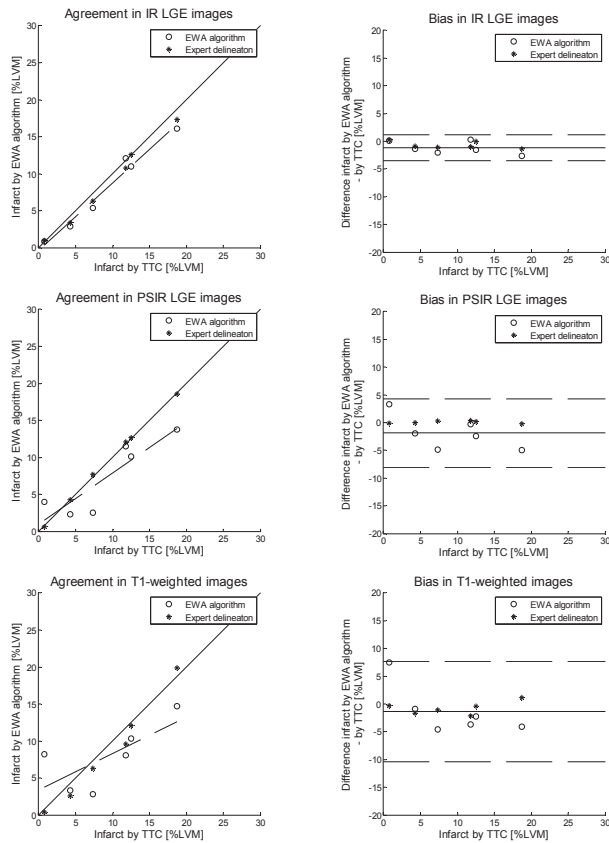
The new automatic algorithm, called EWA, showed a bias to TTC of  $-1 \pm 1$  %LVM ( $R=0.99$ ) in IR images,  $-2 \pm 3$  %LVM ( $R=0.88$ ) in PSIR images and  $-1 \pm 4$  %LVM ( $R=0.75$ ,  $p=0.08$ ) in *ex-vivo* high resolution T1-weighted images (**Figure 16**). Bias to TTC by reference delineation was  $-1 \pm 1$  %LVM ( $R=0.999$ ) in IR images,  $0 \pm 0$  %LVM ( $R=0.994$ ) in PSIR images and  $-1 \pm 1$  %LVM ( $R=0.99$ ) in *ex-vivo* high resolution T1-weighted images (**Figure 16**).

Infarct size by the EWA algorithm was  $15 \pm 8$  %LVM in IR images ( $n=124$ ) with a bias of  $-2 \pm 6$  %LVM ( $R=0.81$ ) compared to the expert delineation. In patients with paired IR and PSIR images ( $n=49$ ) infarct size by the EWA algorithm was  $17 \pm 10$  %LVM in both IR and PSIR images with a bias of  $-1 \pm 5$  %LVM ( $R=0.89$ ) in both IR and PSIR images (**Figure 17**). Bias and variability to reference delineation was lower and regional agreement higher for the EWA algorithm than for the original weighted algorithm and thresholds by EM, 2SD, 3SD and 5SD from remote, FWHM from minimum intensity [22] and FWHM from mean intensity in remote [25] and Otsu's threshold [23] for segmentation in both magnitude and phase sensitive reconstruction (**Table 2**, **Figure 18**).

The EWA algorithm was validated against TTC with bias similar to FWHM from minimum intensity as suggested by Amado et al. [22], the FACT algorithm by Hsu et al. [25] and the original weighted algorithm by Heiberg et al. [26]. However, in our multi-center, multi-vendor patient data a larger variability was found for the original weighted algorithm, and thresholds by n-SD from remote and FWHM from minimum intensity compared to the original experimental validation [22, 26] and single center validation [26], highlighting the importance of validation in representative data such as in multi-center, multi-vendor patient studies. The new automatic algorithm was shown to have a low bias in both experimental validation and multi-center, multi-vendor patients studies for both magnitude and phase sensitive IR LGE images and in T1-weighted high resolution images which showed the capability of intensity classification by EM to adapt to different contrast and image quality.

A

## Validation against TTC



B

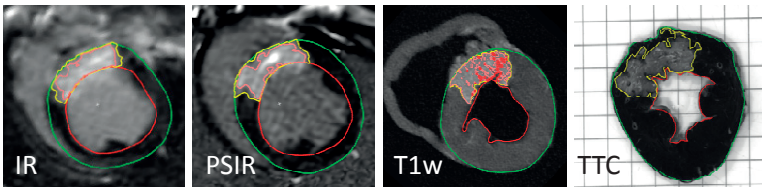
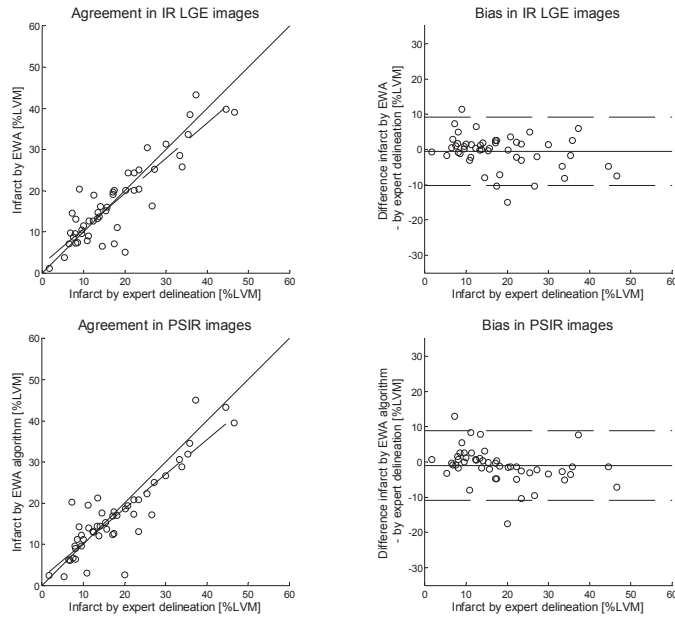


Figure 16 Validation of the EWA algorithm against TTC

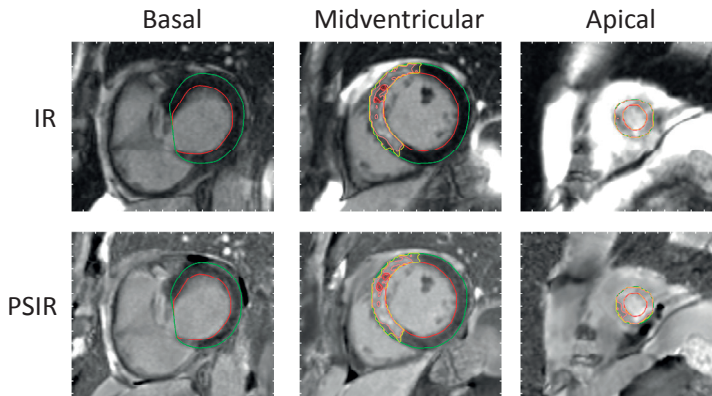
A) Scatter plots (left column) and Bland-Altman plots (right column) of infarct size expressed as LVM for the EWA algorithm against infarct size by TTC in pigs with myocardial infarction imaged after seven days ( $n=6$ ) with IR images (top row), PSIR images (middle row) and *ex-vivo* high resolution T1-weighted images (T1w, bottom row). Left column: solid line = line of identity; dashed line = regression line. Right column: solid line = mean bias; dashed line = mean  $\pm$  two standard deviations.

B) Infarct segmentation by the EWA algorithm in one pig shown in IR LGE, PSIR LGE, high resolution T1-weighted and corresponding TTC-stained slice. Infarct segmentation by the EWA algorithm and by manual delineation in TTC images is shown in yellow. For the automatic EWA segmentation the core of the infarct is shown in pink and microvascular obstruction is shown as the red line within the infarct. Endocardium is delineated in red and epicardium in green.

## A Applicability in paired IR and PSIR LGE images from patients



## B



**Figure 17 Applicability of the EWA algorithm in paired IR and PSIR LGE images from patients in multi-center, multi-vendor studies:**

A) Scatter plots (left column) and Bland-Altman plots (right column) of infarct size expressed as % LVM for the EWA algorithm against infarct size by expert delineation in 49 patients from multi-center studies with paired IR (top row) and PSIR LGE images (bottom row). Left column: solid line = line of identity; dashed line = regression line. Right column: solid line = mean bias; dashed line = mean  $\pm$  two standard deviations.

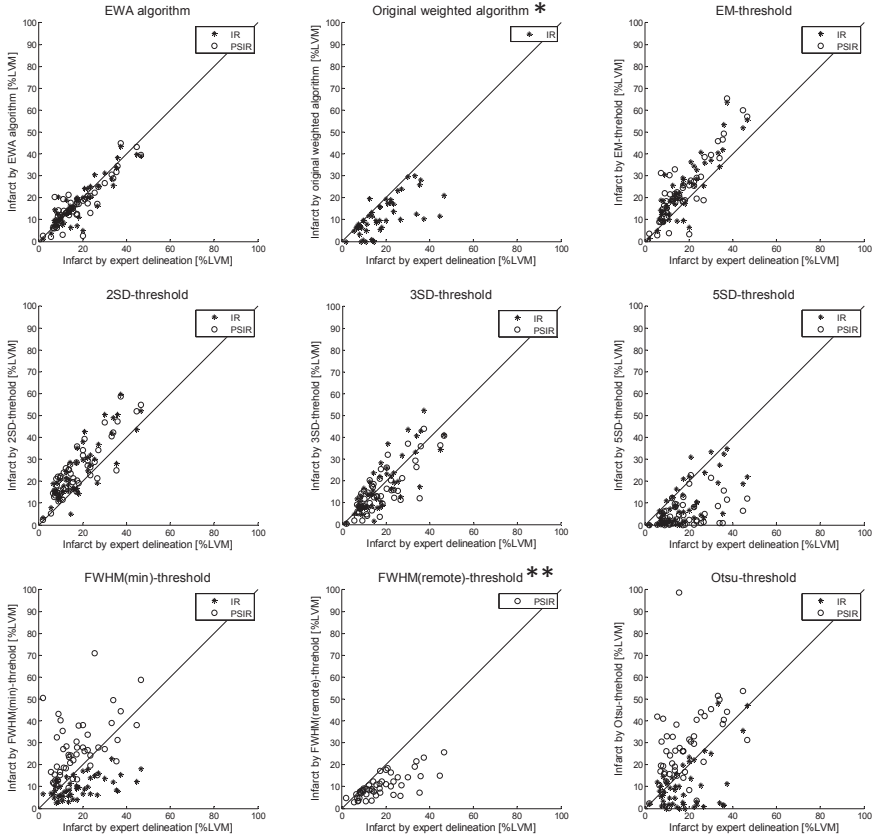
B) Typical segmentation by the EWA algorithm in one patient with paired IR (top row) and PSIR images (bottom row). The automatic EWA segmentation of the infarct is shown in yellow, the core of the infarct is shown in pink and microvascular obstruction is shown as the red line within the infarct. Endocardium is delineated in red and epicardium in green.

	Magnitude IR		Phase sensitive IR	
	Bias to reference delineation [%LVM]	DSC core extent	Bias to reference delineation [%LVM]	DSC core extent
EWA algorithm	-3 ± 5	0.81 ± 0.15	-1 ± 5	0.79 ± 0.15
Original weighted algorithm	-7 ± 8	0.67 ± 0.32	*	*
EM-threshold	6 ± 7	0.67 ± 0.14	6 ± 8	0.68 ± 0.14
2SD-threshold	7 ± 7	0.69 ± 0.15	8 ± 6	0.70 ± 0.13
3SD-threshold	0 ± 7	0.70 ± 0.21	-2 ± 7	0.70 ± 0.19
5SD-threshold	-8 ± 8	0.50 ± 0.33	-13 ± 10	0.36 ± 0.31
FWHM(min)-threshold	-8 ± 9	0.58 ± 0.20	9 ± 12	0.69 ± 0.17
FWHM(remote)-threshold	**	**	-8 ± 7	0.66 ± 0.19
Otsu-threshold	-8 ± 11	0.50 ± 0.35	10 ± 15	0.64 ± 0.20

**Table 2 Bias and regional agreement in paired IR and PSIR LGE images from multi-center patient studies:**

Bias as % of left ventricular mass (%LVM), and regional agreement by DSC to expert delineation for the EWA algorithm, the original weighted algorithm and the threshold method of EM, 2SD, 3SD and 5SD from remote, and FWHM from minimum intensity, FWHM from mean intensity in remote and Otsu's threshold in paired magnitude inversion recovery (IR) and phase sensitive inversion recovery (PSIR) images (n=49). \* the original weighted algorithm was developed for IR images and therefore only applied in such. \*\* the FWHM remote threshold was developed for PSIR images as and therefore only applied in such.

## Applicability in paired IR and PSIR LGE images compared to previously suggested methods



**Figure 18 Applicability in paired IR and PSIR LGE images from multi-center patient studies compared to previously suggested methods for MI quantification:**

Scatter plots of infarct size expressed as % of left ventricular mass (% LVM) against infarct size by expert delineation in 49 patients for the EWA algorithm, the original weighted algorithm and the threshold method of Expectation Maximization (EM), (top row), 2SD, 3SD and 5SD from remote (middle row), and FWHM from minimum intensity, FWHM from mean intensity in remote and Otsu's threshold (bottom row) in paired magnitude inversion recovery (IR) and phase sensitive inversion recovery (PSIR) LGE images. Solid lines = line of identity. \* the original weighted algorithm was developed for IR images and therefore only applied in such.

\*\* the FWHM remote threshold was developed for PSIR and therefore only applied in such.





# Conclusions

In this thesis four new algorithms has been developed for automatic segmentation in CMR; automatic segmentation of the left ventricle in cine SSFP, myocardium at risk in T2-weighted images, myocardium at risk in contrast enhanced SSFP and myocardial infarction in IR and PSIR LGE images. All four algorithms were implemented in the freely available software Segment and all four algorithms were validated against reference delineation with a low bias and high regional agreement to reference delineation.

The major conclusions of each study were

- I. Segment is a well-validated and comprehensive software for cardiovascular image analysis that has been used in a wide range of peer reviewed scientific publications. The software is freely available for research purposes provided that relevant original research related to the software are cited.
- II. The new automatic algorithm for segmentation of the left ventricle was validated in SSFP images from 49 subjects with bias comparable to inter observer variability of manual delineation. With a low level of user input and a regional agreement similar to previous methods the new automatic algorithm is one step closer to automatic segmentation applicable for clinical standard.
- III. The automatic algorithm for myocardium at risk in T2-weighted images was validated in 47 patients with a low bias and high regional agreement to manual delineation. Bias and variability was comparable to inter observer variability of manual delineation and the new algorithm was shown superior to the threshold methods of two standard deviations from remote, full width half maximum and Otsu with regards to both quantitative and regional agreement.
- IV. The automatic algorithm for myocardium at risk in CE-SSFP was validated against manual delineation in 183 patients from multi-center, multi-vendor studies and against SPECT in 16 patients. The new algorithm was shown to have a low bias to manual delineation and SPECT. Bias and variability was comparable to inter observer variability of manual delineation and the new algorithm was shown superior to the threshold methods of two standard deviations from remote, full width half maximum and Otsu with regards to quantitative and regional agreement.
- V. The new automatic algorithm, EWA, for segmentation of myocardial infarction was shown superior to previous threshold methods of standard deviations from remote, full width half maximum and Otsu. The EWA algorithm performed well for both magnitude and phase sensitive LGE images when validated both in

animal studies and multi-center, multi-vendor patient data. Thus, using EM and a weighted approach, as with the EWA algorithm, may serve as a candidate for a clinical standard in quantifying myocardial infarction.

Overall conclusions drawn from the papers included in this thesis are 1) combining intensity classification by EM with *a priori* information was shown superior to previous methods and specifically to threshold methods commonly used, 2) it is important to assess both bias in clinical parameters and regional agreement in validation of new automatic algorithms and 3), it is important to use representative data for validation as in multi-center, multi-vendor patient data.

# Bibliography

1. Grothues, F., et al., Comparison of interstudy reproducibility of cardiovascular magnetic resonance with two-dimensional echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy. *Am J Cardiol*, 2002. **90**(1): p. 29-34.
2. Kim, R.J., et al., Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. *Circulation*, 1999. **100**(19): p. 1992-2002.
3. Fieno, D.S., et al., Contrast-enhanced magnetic resonance imaging of myocardium at risk: distinction between reversible and irreversible injury throughout infarct healing. *J Am Coll Cardiol*, 2000. **36**(6): p. 1985-91.
4. Carlsson, M., et al., Myocardium at risk after acute infarction in humans on cardiac magnetic resonance: quantitative assessment during follow-up and validation with single-photon emission computed tomography. *JACC Cardiovasc Imaging*, 2009. **2**(5): p. 569-76.
5. Sorensson, P., et al., Assessment of myocardium at risk with contrast enhanced steady-state free precession cine cardiovascular magnetic resonance compared to single-photon emission computed tomography. *J Cardiovasc Magn Reson*, 2010. **12**(1): p. 25.
6. Suinesiaputra, A., et al., Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours. *J Cardiovasc Magn Reson*, 2015. **17**(1): p. 63.
7. Schulz-Menger, J., et al., Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) board of trustees task force on standardized post processing. *J Cardiovasc Magn Reson*, 2013. **15**: p. 35.
8. Frangi, A.F., W.J. Niessen, and M.A. Viergever, Three-dimensional modeling for functional analysis of cardiac images: a review. *IEEE Trans Med Imaging*, 2001. **20**(1): p. 2-25.
9. Petitjean, C. and J.N. Dacher, A review of segmentation methods in short axis cardiac MR images. *Med Image Anal*, 2011. **15**(2): p. 169-84.
10. Carlsson, M., et al., Atrioventricular plane displacement is the major contributor to left ventricular pumping in healthy adults, athletes, and patients with dilated cardiomyopathy. *Am J Physiol Heart Circ Physiol*, 2007. **292**(3): p. H1452-9.

11. Jolly, M.P., et al., Combining registration and minimum surfaces for the segmentation of the left ventricle in cardiac cine MR images. *Med Image Comput Assist Interv*, 2009. **12**(Pt 2): p. 910-8.
12. Hu, H., et al., Hybrid segmentation of left ventricle in cardiac MRI using Gaussian-mixture model and region restricted dynamic programming. *Magn Reson Imaging*, 2013. **31**(4): p. 575-84.
13. Codella, N.C., et al., Left ventricle: automated segmentation by using myocardial effusion threshold reduction and intravoxel computation at MR imaging. *Radiology*, 2008. **248**(3): p. 1004-12.
14. Friedrich, M.G., et al., The salvaged area at risk in reperfused acute myocardial infarction as visualized by cardiovascular magnetic resonance. *J Am Coll Cardiol*, 2008. **51**(16): p. 1581-7.
15. Wright, J., et al., Quantification of myocardial area at risk with T2-weighted CMR: comparison with contrast-enhanced CMR and coronary angiography. *JACC Cardiovasc Imaging*, 2009. **2**(7): p. 825-31.
16. Tilak, G.S., et al., In vivo T2-weighted magnetic resonance imaging can accurately determine the ischemic area at risk for 2-day-old nonreperfused myocardial infarction. *Invest Radiol*, 2008. **43**(1): p. 7-15.
17. Burchell, T., et al., Comparing analysis methods for quantification of myocardial oedema in patients following reperfused ST-elevation MI. *Journal of Cardiovascular Magnetic Resonance*, 2011. **13**(Suppl 1): p. M11.
18. McAlindon, E., et al., Quantification of infarct size and myocardium at risk: evaluation of different techniques and its implications. *Eur Heart J Cardiovasc Imaging*, 2015.
19. Johnstone, R.I., et al., Assessment of tissue edema in patients with acute myocardial infarction by computer-assisted quantification of triple inversion recovery prepared MRI of the myocardium. *Magn Reson Med*, 2011.
20. Gao, H., et al., Highly automatic quantification of myocardial oedema in patients with acute myocardial infarction using bright blood T2-weighted CMR. *J Cardiovasc Magn Reson*, 2013. **15**: p. 28.
21. Nordlund, D., et al., Performance of contrast enhanced SSFP and T2-weighted imaging for determining myocardium at risk in a multi-vendor, multi-center setting-data from the MITOCARE and CHILL-MI trials. *Journal of Cardiovascular Magnetic Resonance*, 2015. **17**(Suppl 1): p. P194.
22. Amado, L.C., et al., Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. *J Am Coll Cardiol*, 2004. **44**(12): p. 2383-9.
23. Otsu, N., Threshold Selection Method from Gray-Level Histograms. *Ieee Transactions on Systems Man and Cybernetics*, 1979. **9**(1): p. 62-66.
24. Heiberg, E., et al. Time resolved three-dimensional automated segmentation of the left ventricle. in *Computers in Cardiology*, 2005. 2005.
25. Hsu, L.Y., et al., Quantitative myocardial infarction on delayed enhancement MRI. Part I: Animal validation of an automated feature analysis and combined thresholding infarct sizing algorithm. *J Magn Reson Imaging*, 2006. **23**(3): p. 298-308.

26. Heiberg, E., et al., Automated quantification of myocardial infarction from MR images by accounting for partial volume effects: animal, phantom, and human study. *Radiology*, 2008. **246**(2): p. 581-8.
27. Bondarenko, O., et al., Standardizing the definition of hyperenhancement in the quantitative assessment of infarct size and myocardial viability using delayed contrast-enhanced CMR. *J Cardiovasc Magn Reson*, 2005. **7**(2): p. 481-5.
28. Hsu, L.Y., et al., Quantitative myocardial infarction on delayed enhancement MRI. Part II: Clinical application of an automated feature analysis and combined thresholding infarct sizing algorithm. *J Magn Reson Imaging*, 2006. **23**(3): p. 309-14.
29. Flett, A.S., et al., Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance. *JACC Cardiovasc Imaging*, 2011. **4**(2): p. 150-6.
30. Stirrat, J., et al., Influence of phase correction of late gadolinium enhancement images on scar signal quantification in patients with ischemic and non-ischemic cardiomyopathy. *J Cardiovasc Magn Reson*, 2015. **17**(1): p. 66.
31. Reimer, K.A., et al., The wavefront phenomenon of ischemic cell death. 1. Myocardial infarct size vs duration of coronary occlusion in dogs. *Circulation*, 1977. **56**(5): p. 786-94.
32. Beek, A.M., R. Nijveldt, and A.C. van Rossum, Intramyocardial hemorrhage and microvascular obstruction after primary percutaneous coronary intervention. *Int J Cardiovasc Imaging*, 2010. **26**(1): p. 49-55.
33. Aletras, A.H., et al., Retrospective determination of the area at risk for reperfused acute myocardial infarction with T2-weighted cardiac magnetic resonance imaging: histopathological and displacement encoding with stimulated echoes (DENSE) functional validations. *Circulation*, 2006. **113**(15): p. 1865-70.
34. Simonetti, O.P., et al., An improved MR imaging technique for the visualization of myocardial infarction. *Radiology*, 2001. **218**(1): p. 215-23.
35. Garcia-Dorado, D., et al., Analysis of myocardial oedema by magnetic resonance imaging early after coronary artery occlusion with or without reperfusion. *Cardiovasc Res*, 1993. **27**(8): p. 1462-9.
36. Ubachs, J.F., et al., Myocardium at risk by magnetic resonance imaging: head-to-head comparison of T2-weighted imaging and contrast-enhanced steady-state free precession. *Eur Heart J Cardiovasc Imaging*, 2012. **13**(12): p. 1008-15.
37. Dempster, A.P., N.M. Laird, and D.B. Rubin, Maximum Likelihood from Incomplete Data Via Em Algorithm. *Journal of the Royal Statistical Society Series B-Methodological*, 1977. **39**(1): p. 1-38.
38. Wei, D., et al., A comprehensive 3-D framework for automatic quantification of late gadolinium enhanced cardiac magnetic resonance images. *IEEE Trans Biomed Eng*, 2013. **60**(6): p. 1499-508.
39. Kass, M., A. Witkin, and D. Terzopoulos, Snakes: Active contour models. *International Journal of Computer Vision*, 1988. **1**(4): p. 321-331.
40. Osher, S. and J.A. Sethian, Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 1988. **79**(1): p. 12-49.

41. Cootes, T.F., et al., Information processing in medical imaging Use of active shape models for locating structures in medical images. *Image and Vision Computing*, 1994. **12**(6): p. 355-365.
42. Suinesiaputra, A., et al., Left Ventricular Segmentation Challenge from Cardiac MRI: A Collation Study, in *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, O. Camara, et al., Editors. 2012, Springer Berlin Heidelberg. p. 88-97.
43. Zijdenbos, A.P., et al., Morphometric Analysis of White-Matter Lesions in Mr-Images - Method and Validation. *Ieee Transactions on Medical Imaging*, 1994. **13**(4): p. 716-724.
44. Abdel-Aty, H., et al., Delayed enhancement and T2-weighted cardiovascular magnetic resonance imaging differentiate acute from chronic myocardial infarction. *Circulation*, 2004. **109**(20): p. 2411-6.
45. Khan, J.N., et al., Comparison of semi-automated methods to quantify infarct size and area at risk by cardiovascular magnetic resonance imaging at 1.5T and 3.0T field strengths. *BMC Res Notes*, 2015. **8**(1): p. 1007.
46. Cerqueira, M.D., et al., Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Circulation*, 2002. **105**(4): p. 539-42.
47. Dice, L.R., Measures of the Amount of Ecologic Association between Species. *Ecology*, 1945. **26**(3): p. 297-302.
48. Erlinge, D., et al., Rapid endovascular catheter core cooling combined with cold saline as an adjunct to percutaneous coronary intervention for the treatment of acute myocardial infarction. The CHILL-MI trial: a randomized controlled study of the use of central venous catheter core cooling combined with cold saline as an adjunct to percutaneous coronary intervention for the treatment of acute myocardial infarction. *J Am Coll Cardiol*, 2014. **63**(18): p. 1857-65.
49. Atar, D., et al., Effect of intravenous TRO40303 as an adjunct to primary percutaneous coronary intervention for acute ST-elevation myocardial infarction: MITOCARE study results. *Eur Heart J*, 2015. **36**(2): p. 112-9.
50. Ugander, M., et al., Quantification of myocardium at risk in myocardial perfusion SPECT by co-registration and fusion with delayed contrast-enhanced magnetic resonance imaging--an experimental ex vivo study. *Clin Physiol Funct Imaging*, 2012. **32**(1): p. 33-8.
51. Soneson, H., et al., An improved method for automatic segmentation of the left ventricle in myocardial perfusion SPECT. *J Nucl Med*, 2009. **50**(2): p. 205-13.
52. Jablonowski, R., et al., Contrast-enhanced CMR overestimates myocardial infarction size on day 1 but not day 7 relative to TTC in a swine model: Mechanistic insights using extracellular volume measurements. *Journal of American College of Cardiology Cardiovascular Imaging*, 2015. **In Press**.
53. Hansson, M.J., et al., Differences in the profile of protection afforded by TRO40303 and mild hypothermia in models of cardiac ischemia/reperfusion injury. *Eur J Pharmacol*, 2015. **760**: p. 7-19.







**SOFTWARE**

**Open Access**

# Design and validation of Segment - freely available software for cardiovascular image analysis

Einar Heiberg\*, Jane Sjögren, Martin Ugander, Marcus Carlsson, Henrik Engblom, Håkan Arheden

## Abstract

**Background:** Commercially available software for cardiovascular image analysis often has limited functionality and frequently lacks the careful validation that is required for clinical studies. We have already implemented a cardiovascular image analysis software package and released it as freeware for the research community. However, it was distributed as a stand-alone application and other researchers could not extend it by writing their own custom image analysis algorithms. We believe that the work required to make a clinically applicable prototype can be reduced by making the software extensible, so that researchers can develop their own modules or improvements. Such an initiative might then serve as a bridge between image analysis research and cardiovascular research. The aim of this article is therefore to present the design and validation of a cardiovascular image analysis software package (Segment) and to announce its release in a source code format.

**Results:** Segment can be used for image analysis in magnetic resonance imaging (MRI), computed tomography (CT), single photon emission computed tomography (SPECT) and positron emission tomography (PET). Some of its main features include loading of DICOM images from all major scanner vendors, simultaneous display of multiple image stacks and plane intersections, automated segmentation of the left ventricle, quantification of MRI flow, tools for manual and general object segmentation, quantitative regional wall motion analysis, myocardial viability analysis and image fusion tools. Here we present an overview of the validation results and validation procedures for the functionality of the software. We describe a technique to ensure continued accuracy and validity of the software by implementing and using a test script that tests the functionality of the software and validates the output. The software has been made freely available for research purposes in a source code format on the project home page <http://segment.heiberg.se>.

**Conclusions:** Segment is a well-validated comprehensive software package for cardiovascular image analysis. It is freely available for research purposes provided that relevant original research publications related to the software are cited.

## Background

Applied medical research is becoming more and more dependent on imaging for evaluation of the therapeutic effects of new drugs or therapies. Thus, dedicated image analysis software is needed for quantitative medical imaging. Commercially available software often offers limited functionality and frequently lacks the validation that is required for clinical studies. On the other hand, open source software offers transparency and the ability to

modify the source code is well-suited to academic research since it gives researchers the ability to see exactly how the algorithms are implemented. The current trend among research grant organisations is that the results of government-funded projects should be published in open access journals or should be otherwise publicly available. Consequently, open access publishing has had a noticeable effect on the ease with which scientific results become available.

Publication of results in peer-reviewed journals is the traditional way of documenting and mediating progress in science. For classical sciences such as medicine and

\* Correspondence: [einar@heiberg.se](mailto:einar@heiberg.se)

Department of Clinical Physiology, Lund University and Lund University Hospital, Lund, Sweden

physics, this is normally sufficient since the reader can incorporate the information given in the publication into his or her own research. However, medical image analysis research often involves complex algorithms, and one cannot easily incorporate the results into one's own research since the algorithms described in the scientific papers usually need to be re-implemented for use by other research groups [1]. Thus, we believe that releasing medical image analysis software to other research groups as freeware has the potential to have a profound effect on medical image analysis and applied medical imaging research. If the source code is reusable and written in a standardized way, the application can be modified or extended with new algorithms. This would permit the development of new algorithms or refinement of existing algorithms in order to satisfy requirements that arise in a clinical research setting. Furthermore, scripting capabilities for medical image analysis software may open up new lines of research that were previously untestable, since manual analysis would not have been feasible. For example, in a recent study it was possible to classify 72 regional myocardial sectors according to the neighbouring sectors and to track them in 22 patients over 5 points in time, thus generating over 50,000 classified data points [2]. This task would have been impossible without advanced scripting capabilities in the analysis software.

The Segment cardiac image analysis software package, which is the subject of this article, was originally developed by the first author and was released in 2005 on a freely available basis. Since then it has been downloaded by more than 2,000 unique users in 74 countries, and approximately 300 research groups. To date, it has been referenced in more than 40 scientific publications. One advantage of freely available medical image analysis software is that it facilitates multi-centre clinical trials since all the participating sites can use the same software. The authors are aware of two ongoing multi-national and multi-centre studies that are using Segment software. Until recently, the software was distributed as a precompiled Windows application. If the software was available in a source code format, other researchers could contribute with their own modules and improvements. Image processing experts could then directly integrate new features into the software and make these improved algorithms available to other researchers. In this way, algorithm developers would be able to focus on algorithm development and make use of the common basic functionality of the software such as image loading, image display, user interactions etc. We believe that this may not only increase scientific productivity, but more importantly it may also provide a bridge between the very latest image processing ideas and applied clinical research. Today, many very promising image processing

ideas never reach clinical routine because it is too cumbersome to write prototype software that is sufficiently user-friendly to be used by clinical researchers. Thus, the aim of this article is to present the design and validation of a cardiovascular image analysis software package and to announce its release in a source code format.

## Implementation

What follows is an outline of some of the important details concerning design and implementation of the software.

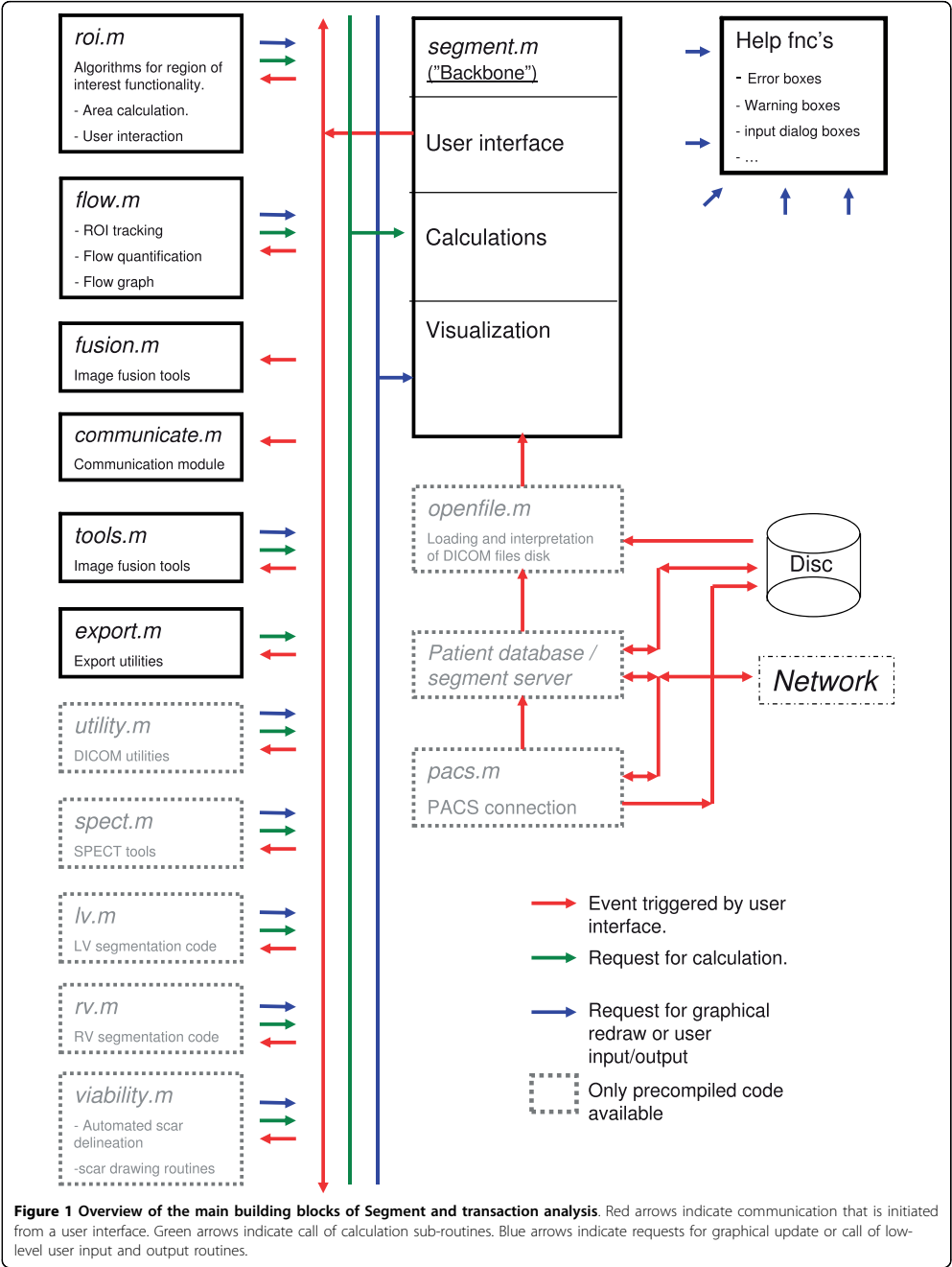
### A. Overview

The software can be divided into 14 main functional blocks. An overview of how these building blocks communicate and relate to each other is given in Figure 1. Each block is implemented as a separate Matlab file and documented in detail in the technical manual. Closer descriptions of each building block are beyond the scope of this article. The grey blocks are not included in the source code, but will be accessible as precompiled code.

An overview of the image processing algorithms is given in Table 1, an overview of manual image processing tools is given in Table 2, and finally the export, import and reporting capabilities are described in Table 3. For each algorithm or functionality, appropriate references to publications where the functionality has previously been described, if applicable.

Each loaded image or image stack (a single image (2D), a time-resolved single image (2D+T), a stack of images covering a multi-slice image volume (3D) or a time-resolved multi-slice image volume (3D+T)) is stored as a struct, and has fields for storing contours, image data, image orientation, resolution, image acquisition details, delineations, annotations and measurements. Handles to the graphical user interface and some temporary data are stored in a global data structure. A set of low-level input and output user interface routines was implemented to improve portability of the code. An object-oriented system for graphical user interfaces is employed to improve the ability to maintain the code and simplify development of new user interfaces.

The software was designed with clinical research in mind, and to maximise work flow and user-friendliness. A screen shot of the main graphical user interface is shown in Figure 2. The internal file format used by Segment was designed so that a complete patient examination can be stored together with all measurements and annotations. This allows the clinical researcher to go back and see how the delineations were made, which is often not possible with commercially available software packages. The internal file format also allows the user to batch process multiple data sets and to export quantitative data in a spreadsheet format.



**Table 1 Automated or semi-automated image processing tools in Segment**

Algorithm	Dimensionality	Reference	Section
<i>Ventricle segmentation</i>			
- Left ventricle	2D, 2D+T, 3D, 3D+T	[5]	C.
- Semi-automatic tools for right ventricle	2D, 2D+T, 3D, 3D+T	*	C.
<i>Flow</i>			
- Phase unwrapping algorithm	2D+T, 3D+T, 3+3D+T	*	D.
- Phase background correction	2D+T, 3D+T, 3+3D+T	*	D.
- Automated vessel tracking	2D+T, 3D+T	*	D.
- Flow visualization	2D+T	*	D.
<i>Delayed enhancement/viability</i>			
- Quantification of infarct size	3D	[3,4]	H.
- Infarct extent	3D	[11,22]	H.
<i>General Object Segmentation</i>			
- Fast levelset	3D, 2D+T, 3D+T	[25]	I.
	3D, 2D+T, 3D+T	*	I.
<i>SPECT</i>			
- Left ventricle segmentation	3D	[28]	J.
- Defect size	3D	#	J.
- Gated SPECT segmentation	3D+T	#	J.

Dimensionality: 2D) works on two dimensional images, 2D+T) works on time resolved two dimensional images, 3D) works on three dimensional images, 3D+T) works on time resolved three dimensional images, 3+3D+T) works on three component three dimensional time resolved images. Reference: (X) previously published in reference X, #) previously unpublished data, manuscript submitted, \*) algorithm presented for the first time in this study. Section: Refers to the Result section where functionality is described.

**Table 2 Manual image processing tools in Segment**

Algorithm	Dimensionality	Ref	Section
<i>Image visualization tools</i>			
- Contrast adjust + auto contrast	2D, 2D+T, 3D, 3D+T, 3+3D+T	*	B.
- Multi view/panel support	2D, 2D+T, 3D, 3D+T, 3+3D+T	*	B.
- Image plane intersection	3D, 3D+T, 3+3D+T	*	B.
<i>Manual contouring tools</i>			
	2D, 2D+T, 3D, 3D+T	*	E.
<i>Region of interest analysis (ROI)</i>			
- Signal intensity quantification	2D, 2D+T, 3D+T	*	F.
- Histogram analysis	2D, 2D+T, 3D+T	*	F.
- Visual ROI analysis	2D, 2D+T	*	F.
- Area tools	2D, 2D+T	*	F.
- Volume tools	3D, 3D+T	*	F.
<i>Linear measurements</i>			
	2D, 2D+T	*	G.
<i>Annotation points</i>			
	2D, 2D+T, 3D, 3D+T	*	G.
<i>Image fusion</i>			
	3D	[27]	K.
<i>Reformatting image tools</i>			
- Multi planar reconstruction	3D, 3D+T	*	L.
- Resampling	2D, 2D+T, 3D, 3D+T	*	L.

Dimensionality: 2D) works on two dimensional images, 2D+T) works on time resolved two dimensional images, 3D) works on three dimensional images, 3D+T) works on time resolved three dimensional images, 3+3D+T) works on three component three dimensional time resolved images. Reference: (X) previously published in reference X, #) previously unpublished data, manuscript submitted, \*) algorithm presented for the first time in this study. Section: Refers to the Result section where functionality is described.

**Table 3 Export, import and reporting capabilities of Segment**

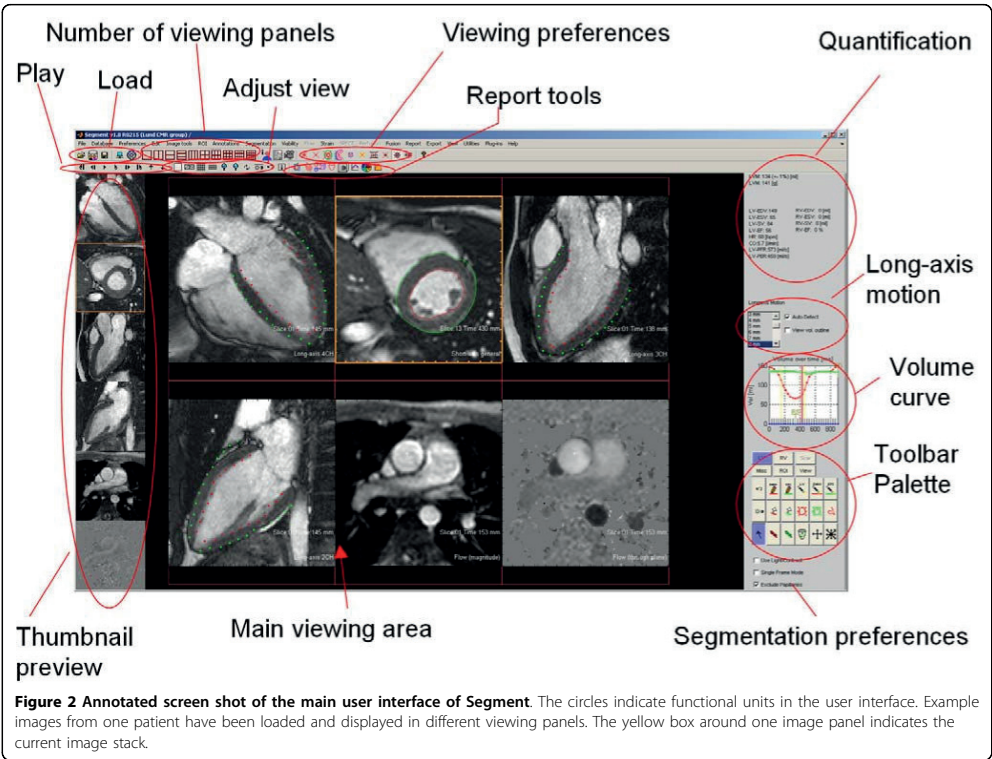
Algorithm	Dimensionality	Reference	Section
DICOM import and manipulation	2D, 2D+T, 3D, 3D+T	*	A
Movie recording capacity	-	*	B
Wall thickening analysis	2D+T, 3D+T	*	C
Polar plot of function and infarct	3D, 3D+T	[29]	M
Batch export to statistical software	-	*	N
Communication module to facilitate multicenter trials	-	*	O
Plug-in capabilities	-	*	P

Dimensionality: 2D) works on two dimensional images, 2D+T) works on time resolved two dimensional images, 3D) works on three dimensional images, 3D+T) works on time resolved three dimensional images, 3+3D+T) works on three component three dimensional time resolved images, -) dimensionality not applicable. Reference: [X]) previously published in reference X, #) previously unpublished data, manuscript submitted, \*) algorithm presented for the first time in this study. Section: Refers to the Result section where functionality is described.

**B. Programming environment**

The software package is written in Matlab and, for time-critical sections of the code, standard ANSI-C is used coded with Mex-wrappers so that they can be called from Matlab. The complete software project consists of about 90,000 lines of Matlab code and about 10,000 lines of C code. In total, there are 44 separate user interface panels. Proper version control software is

necessary when managing a programming project of this size. Version control is managed using the open source solution Subversion (SVN, <http://subversion.tigris.org>) with Tortoise SVN <http://tortoisesvn.tigris.org> as a shell extension. Feature requests and bug reports are managed using the web-based open source software Trac <http://trac.edgewall.org>. Up until now, the whole application has been compiled as a stand-alone application



**Figure 2 Annotated screen shot of the main user interface of Segment.** The circles indicate functional units in the user interface. Example images from one patient have been loaded and displayed in different viewing panels. The yellow box around one image panel indicates the current image stack.

and distributed together with the Matlab Compiler Runtime Environment. Upon publication of this article, the source code, the user manual and the technical manual will be available on the Segment home page <http://segment.heiberg.se>. Precompiled versions of Segment will still be available for Windows and Linux (Ubuntu distribution). Precompiled object files for the C code will be made available for 32-bit and 64-bit Windows operating systems, 32-bit and 64 bit Linux, and Mac OS X.

### C. Stability, accuracy and validity

For clinical image analysis, the following requirements of a software package are crucial: (1) stability, (2) high performance, and (3) accuracy and validity.

To achieve stability, the software was designed so that even when run-time errors occur, the application should not crash and the user interface should not end up in a state in which the user must to restart the application.

To optimise performance, effective memory management and highly optimised routines are essential when designing the software. Cardiac imaging is particularly demanding since the data sets are frequently large and multidimensional with regard to space and time. Great care was taken to avoid duplicating image data unless absolutely necessary. Also, time-critical routines were optimised and coded in C.

To achieve accuracy and validity, validation considerations were incorporated into the design process by placing all calculations of distances, areas, volumes and region of interest mask generation in well-validated subfunctions. Care was taken at the design phase to avoid loss of accuracy due to loss of numerical precision. One such consideration was to represent contours, measurements and regions of interest with double-precision floating point numbers. Surprisingly, this is not often the case in commercially available software tools where pixel-based approaches are frequently used, which can be quite misleading—especially in small regions of interest. Even with a robust underlying design, it remains a challenge to maintain a strict quality policy that allows use for clinical research. The solution that we have chosen is to write an extensive test script that runs on archived test data. The test script output from the software is then compared with known accurate results from previously validated scientific publications. To ensure that the software maintains a high standard, the complete software repository will only be made available to a limited number of trained developers and incorporation of user-contributed code in the code base will only be done after careful testing and quality control. With the use of the test script, it is possible to quickly test the entire software project and uncover unexpected side effects when the code is modified. Coding standards and quality policies are given in the technical manual, which is available on the project home page.

### D. Software maintenance

Medical imaging is developing rapidly and as a result medical image analysis software must be continuously refined and maintained. To ensure long-term maintenance of the code, we have chosen a solution whereby a company was formed to support and commercialise the software for use in clinical practice or by commercial users. Lund University is a shareholder in this company, and researchers have and will continue to have access to the current and future versions of the source code of the software.

### E. Terms of licence

Segment is freely available for academic investigational research use (studies paid by government-derived funds or donations) provided that the original research publications relevant to the software are cited. The software is also free for educational purposes. Note that the license terms do not generally include trials paid by pharmaceutical companies. For commercial use, Segment is sold and supported by the company Medviso AB, Lund, Sweden. Individuals or organisations are not allowed to compile software products derived from Segment that are to be sold commercially or shipped together with other commercial products without written permission from Medviso AB.

## Results

Segment is a full-featured software tool for cardiovascular image analysis and to date, it has been used in a wide range of publications ranging from technical algorithm descriptions [3-5] to applied research on the effects of cardiac gene therapy [6], perfusion MRI [7], perfusion multidetector CT [8], applied human physiology [9], validation of an imaging technique in clinical cardiology [10], analysis of infarction with MRI [11] and MDCT [12], and analysis of microinfarction [7], regional cardiac function [13], for the first time quantitatively determine the infarct evolution in man [14], brain imaging [15] and also experimental imaging in rodents [16]. The software has been developed with a view to its use in cardiovascular magnetic resonance imaging (MRI) and myocardial perfusion single-photon emission computed tomography (SPECT), but in principle it can be used for image analysis in any organ system, and it has also been used for image analysis in computed tomography (CT) and positron emission tomography (PET). The following sections deal with the main features of Segment. Each section presents an overview of validation results and procedures where applicable.

### A. Loading of DICOM images

DICOM images from all major MRI vendors, including both human and animal scanners (Bruker, GE, Philips, Siemens and Varian) can be loaded into the software. Correctness of the loaded image data is validated by the

test script, which loads a large number of image types from all vendors and compares the results to previously manually validated results using ImageJ or proprietary vendor software. Checks are made for correct image sorting, image resolution, time increment between time frames, slice thickness etc.

## B. Image display

A large set of image display tools has been implemented. Examples of functionality include simultaneous display of multiple image stacks, viewing of contours and regions of interest, image stack intersections, adjustment of image contrast and brightness, and scrolling over time and between image slices. Each image stack can have its own colour scale or contrast settings, and 12-bit color mapping is used internally. The user interface was developed to maximise the image display area without compromising user friendliness. Furthermore, multiple monitors are supported to maximise the display area and improve the work flow. Image display tools cannot be evaluated quantitatively, but the functionality is carefully tested by the test script. Tools for recording movies are also incorporated in the software. This greatly facilitates preparation of scientific presentations that include movies.

## C. Automated segmentation of cardiac ventricular dimensions in MRI

Automated segmentation of the left ventricle in MRI was the first image process algorithm implemented in Segment. The algorithm has been described and validated [5,17]. There was an excellent correlation between automated segmentation and manual segmentation for end diastolic volume (EDV),  $R^2 = 0.99$ , with a mean error of  $-1 \pm 11$  ml, and left ventricle mass (LVM),  $R^2 = 0.94$ , with a mean error of  $4 \pm 15$  ml [5]. Tools for semi-automatic delineation of the right ventricle have also been incorporated. From the automatically or semi-automatically segmented surfaces, wall thickness, wall thickening, and fractional wall thickening can be calculated [5,17]. Typical computational time for a standard Windows XP desktop PC (Intel Dual Core 2 GHz, Bus speed 770 MHz and 2 GB RAM) is about 10 seconds for a typical data set with 12 slices, and 30 timeframes.

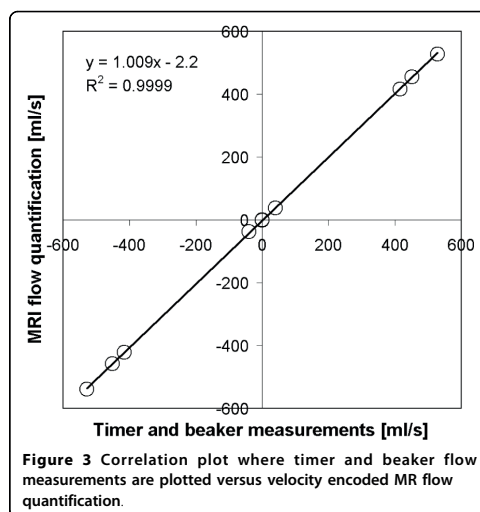
## D. Flow quantification

Flow measurements are of major importance in cardiovascular research, and velocity encoded phase contrast (PC) magnetic resonance imaging (MRI) is the golden standard for the *in vivo* quantification of blood flow in large vessels. It has been shown that modern MRI scanners may have phase offsets due to eddy currents, which can have a large effect on clinical flow measurements [18]. Segment has tools for compensation of linear and higher-order background phase offsets due to eddy currents or Maxwell effects [19]. It is possible to use automatic detection of stationary tissue based on the

temporal standard deviation of the phase or to use manual regions of interest.

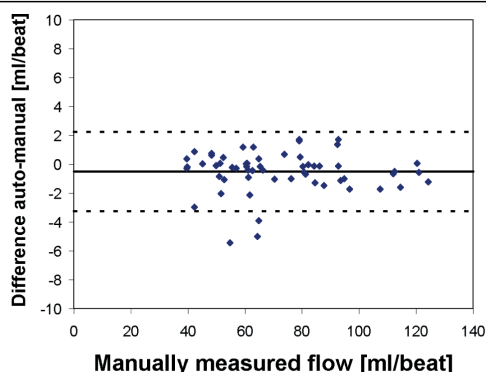
When imaging velocities higher than the chosen velocity encoding range, phase aliasing will occur and cause wrap-around artifacts in the quantitative visualization of velocities. Tools to compensate for such aliasing effects have been implemented and used successfully [20]. The algorithm detects temporal discontinuities in the phase, and pixels with a temporal phase jump pair are unwrapped. The tool also includes a graphical user interface for manual correction.

Validation of flow measurements has been performed both in phantom experiments and in patients. Phantom measurements were performed using gravity-driven flow at 5 different flow rates through a silicon gel with two holes 26 mm in diameter. The true rate of flow was measured by beaker and timer. The agreement between measured flow and beaker and timer was excellent ( $y = 1.009x - 2.2$  ml,  $R^2 = 1.00$ ). Figure 3 shows a correlation plot for this experiment. In patient images, one experienced observer outlined the ascending and descending aorta in 32 patients. In total, 64 regions of interest were analysed both manually and using automated vessel delineation. In the automated vessel delineation, the manually outlined vessel contour from the first time frame was taken as input to the algorithm. In 4 of the 64 vessels (6%), the automated vessel tracking failed due to poor image contrast and imaging artifacts, and resulted in a difference greater than 10 ml and large visual overestimation in vessel area. These vessels were excluded from further analysis. Bias and variability between total net flow for the manual vessel delineation and



**Figure 3** Correlation plot where timer and beaker flow measurements are plotted versus velocity encoded MR flow quantification.





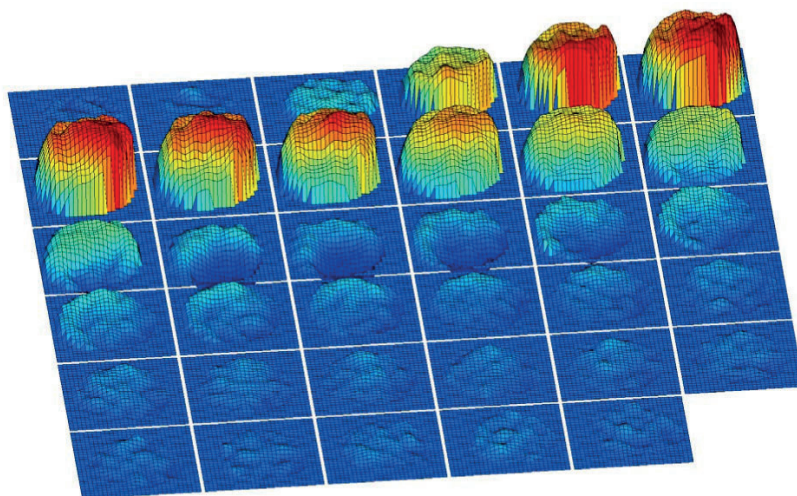
**Figure 4** Difference in total net flow comparing automated and manual vessel delineation. Bias  $\pm 2$  SD is indicated in the plot.

automated vessel delineation was  $-0.5 \pm 2.8$  ml/beat for the remaining 60 vessels. Figure 4 illustrates the difference in total net flow between automated and manual vessel delineation. Total time for automated vessel tracking is about 1.5 seconds on an ordinary desktop PC for a typical data set with 35 time frames. Besides the quantitative tools for flow mapping, there are also visualisation tools for visualisation of flow profiles. Figure 5 illustrates vessel flow profiles over time. The first time

frame is at the top left, and time is increasing along each row.

#### E. Tools for drawing object contours

All the necessary tools for manually drawing object contours, and regions of interest, linear measurements and annotations are implemented. The same tools can be used transparently to correct automated delineated object contours. All drawing tools include full undo capabilities. All quantitative measures from object contours



**Figure 5** Example of vessel flow profile visualisation over time in the human aorta of a healthy volunteer. The first time frame is at the top left and time is increasing along each row. Top right vessel is peak systolic time frame. Note the relative skewedness of the flow in the healthy volunteer.



rely on the same low-level quantification algorithms. These low-level algorithms were validated using computer phantoms implemented in the software, and are included in the test script.

#### F. Region of interest analysis

Tools for analysing regions of interest (ROIs) are implemented. Details about signal intensities and area are available. Areas of the regions of interest are measured by accurate polygon calculations, but signal intensity and signal intensity statistics are calculated by performing statistics on discrete pixels inside the ROI. Basic statistics such as mean intensity and standard deviations over time are available. Typical calculation time for calculating statistics for one ROI in a typical data set with 35 time frames is about 0.2 s on an ordinary Windows XP desktop PC, Intel Dual Core 2 GHz, Bus speed 770 MHz and 2 GB RAM.

Image intensities are stored in Segment as single precision floating-point numbers and internally scaled into the range 0-1. Signal intensity measurements are done with single-precision arithmetics which give relative errors of about  $10^{-6}$ . However, note that the DICOM standard specifies storage of images using between 12 and 16 bits. For CT images, they are internally stored in Segment as 16-bit signed integers and are converted to single-precision floats before any arithmetics besides general object segmentation or display.

Detailed statistics such as distribution are available as histograms of pixel intensities. Measuring signal intensities in small ROIs using discrete pixels introduces a sampling error. This sampling error is quantified in Figure 6 for circular ROIs. For regions of interest of a size between 0 and 1 cm<sup>2</sup>, the corresponding error is  $0.07 \pm 2.5$  for a pixel resolution of 1 mm,  $-0.17 \pm 0.88$  for a pixel resolution of 0.5 mm, and  $-0.1 \pm 0.27$  for a pixel resolution of 0.25 mm. Measurements for maximum values, minimum values, and full-width half maximum values, and minimum and maximum temporal derivatives can also be exported. It is also possible to apply smoothing prior to these calculations. The optional smoothing applied is a user-adjustable Gaussian filtering, and is implemented using Normalized Averaging to account for edge effects at the beginning and end of the signal.

#### G. Linear measurements and annotation points

Linear measurements can be made in single time frame and the distances between start and end points are calculated as the standard Euclidian distance. Distances and positions are stored as double-precision floats, and thus positions and distances can be measured with sub-pixel accuracy. Annotation points, both static in time and time-resolved, can be placed in the image volume, and coordinates for these annotation points can be

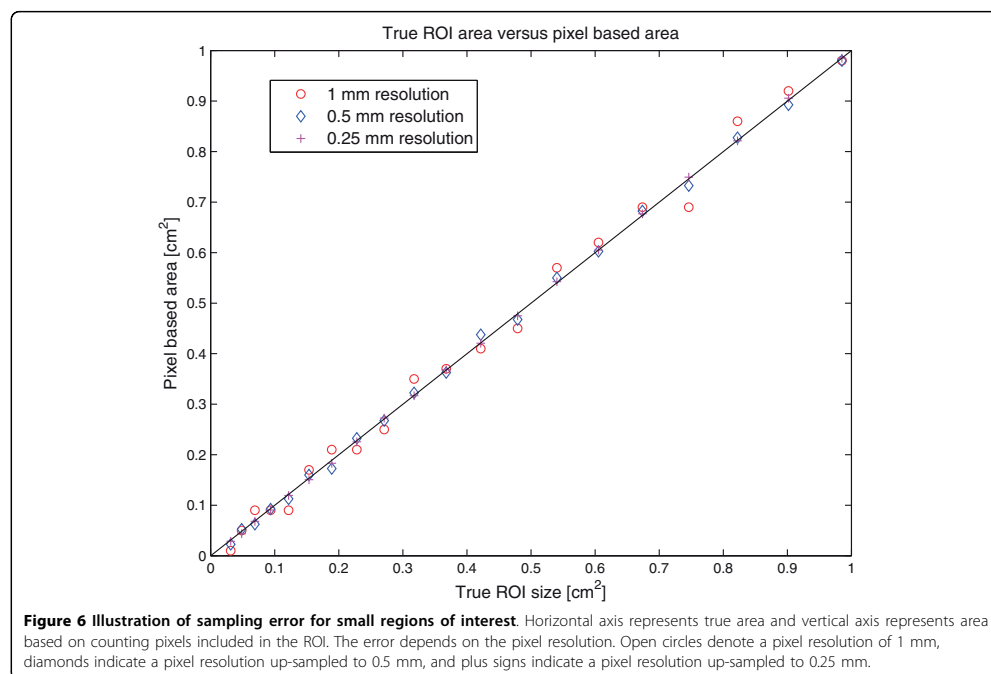
exported. This also allows measurement of distances in time-resolved three-dimensional space.

#### H. MRI viability analysis

Delayed contrast enhancement MRI can be used to differentiate between viable and necrotic or fibrotic tissue, owing to changes in the extracellular distribution volume after irreversible cell injury [21]. Automated tools for delineation of hyper-enhanced regions are implemented and have been clinically validated [3] and used in several studies, such as [22] and [11]. Hyper-enhanced regions can be expressed in ml, as a percentage of left ventricle mass, or as endocardial extent [22]. Furthermore, the method for delineation of hyper-enhanced regions has been improved with a novel paradigm to account for the partial volume effect, where the hyper-enhanced pixels are weighted depending on the image intensity [4]. The method was validated in 8 animals by *in vivo* MRI compared to high-resolution *ex vivo* MRI. Further validation was performed in 40 computer phantoms and 40 patients, by comparing the results to manual delineations from three experienced observers. Values of mean bias  $\pm$  variability (or standard deviation), expressed as a percentage of left ventricular myocardium (%LVM), were  $-0.3 \pm 1.3\%$  (animals),  $-1.2 \pm 1.7\%$  (phantoms) and  $0.3 \pm 2.7\%$  (patients). The new weighted algorithm had lower variability than the previously published approach of dichotomously classifying pixels as wholly infarcted or not ( $2.7$  vs  $7.7\%$  LVM,  $P < 0.01$ ) and was not statistically significantly different from inter-observer variability for bias ( $P = 0.31$ ) or variability ( $P = 0.38$ ). Also, calculation of weighted infarct transmuralities has been implemented and validated [23]. Typical calculation time for automated infarct delineation for a typical data set with 12 slices is about 0.1 s on an ordinary desktop PC.

#### I. General image object segmentation tools

Tools for segmentation of general objects have been implemented, and they are based on an approximate fast-level set algorithm [24]. The segmentation can be exported as a three-dimensional mesh. This method has been described in detail already [25]. Moreover, a novel prototype-based segmentation algorithm has been developed. This algorithm is capable of learning how to segment an object based on few examples, typically 5-10. This can be contrasted to traditional statistical model-based approaches, which generally required 50-100 training cases. In a test case the aorta was delineated in 10 healthy volunteers imaged using a steady-state free precession sequence, resolution  $1.6 \times 1.6 \times 1.0$  mm. Five volunteers were used as a learning set and five as test case. The volumetric error was  $4 \pm 5\%$  when measured as volumetric error and  $0.58 \pm 0.06$  mm when measured as distance error.



#### J. Left ventricle segmentation in SPECT images

Automated delineation of the left ventricle in myocardial perfusion SPECT has been implemented. The result of the fully automated segmentation in myocardial perfusion SPECT images of 100 patients was compared to that from manual planimetry of MRI images of the same patients as a reference standard. The mean error and variability compared to MRI was  $6 \pm 15\%$  LVM, which was significantly lower than for the commercially available algorithm Quantitative Perfusion SPECT (QPS, Cedars-Sinai Health Systems):  $18 \pm 19\%$  LVM [26]. Typical time for fully automated LV segmentation is about 8 s on an ordinary desktop PC.

#### K. Image fusion

A tool for manual rigid body co-registration of image stacks was implemented and was evaluated in an imaging study where *ex vivo* data from 19 pigs using myocardial perfusion SPECT and MRI were successfully co-registered [27].

#### L. Reformatting image tools

General image tools such as multi-planar reconstruction, re-sampling and rotation have been implemented. Accuracy of re-sampling was validated by re-sampling of computer phantoms with known geometric properties. The image quality of re-sampled images is maximized

by using bi-cubic interpolation with anti-aliasing filters. Testing of the general image tools is included in the test script.

#### M. Polar plots

Quantitative results of regional functional parameters such as wall thickening and infarct transmural from viability analysis can be displayed as polar plots, including segmentalisation of the left ventricle according to the 17-segment model endorsed by the American Heart Association [28]. Development and validation of this functionality has been described [29], and an example of how this can be used to register quantitative data in consecutive MRI studies has been published [2].

#### N. Export capabilities

A wide range of exportation algorithms were implemented. Export direct to the system clipboard can be done from most graphical user interface panels. Data in the clipboard can then be pasted directly into statistical spread-sheet software such as Microsoft Excel. Tools for batch exportation from multiple files are also implemented. This avoids tedious manual interaction, and it also eliminates manual mistakes in the exportation process.

#### O. Communication module

A module to facilitate communication between research groups and to facilitate multi-national and multi-centre

studies is included in the software. Since the software is freely available to researchers, all participating sites can use the same software. Images are loaded into the software by the imaging site. Using the software, all images from one CMR examination are (optionally) analysed, anonymised, compressed and electronically transferred to a central server set up by the coordinating centre.

The coordinating centre can control the total image flow by distributing "site keys" to each participating site. The site key includes instructions to the imaging centre and tells the software where images from a specific site should be stored. No passwords or usernames need to be sent out to participating sites. The site key also contains the encryption/decryption key.

#### **P. Plug-in and scripting capability**

Possibilities for writing one's own dedicated plug-ins have been implemented and details are given in the technical manual. A detailed description is outside the scope of this article. A plug-in template is available, consisting of less than 30 lines of Matlab code, and user plug-ins are available directly main menu in Segment. Users have access to the complete internal data structure and have control over all elements of the graphical user interface. Scripting can be performed directly in Matlab, thus providing powerful scripting possibilities.

#### **Q. Test script**

Testing includes user interface, display, and low-level algorithms such as automated LV segmentation, vessel delineation, flow quantification, DE-MRI viability analysis etc. Currently, the test script comprises 66 main test cases and a total of 316 tests. The test script is continuously refined as new features are added to the program. A script that checks for broken call-back links in the user interface is also implemented.

### **Discussion**

Previous work on extendable software for medical imaging can be divided into two major categories: (1) complete applications, and (2) toolkits such as MITK <http://www.mitk.org> [1]. These two different approaches to extendable medical analysis software have their own strengths and weaknesses. Early on, we considered whether Segment should be a toolkit or a complete application. It was decided that an application approach would be most beneficial to clinical research since it will allow the great majority of users to be able to use it without having to write any source code. One pioneering and freely available extendable medical imaging software package is ImageJ <http://rsbweb.nih.gov/ij/>, written by Rasband at the U.S. National Institutes of Health. ImageJ has become widely popular, probably because of the open platform that allows users to write their own plug-ins, thus rapidly increasing the applicability of the software. Another open source general viewing and

image processing application that also has a complete plug-in architecture is OsiriX <http://www.osirix-viewer.com/>, which runs under Mac OS X. Yet another well-known application is Slicer <http://www.slicer.org/> [30]. One example of extendable software for medical imaging that is not open source is Analyze (AnalyzeDirect, Lenexa, KS; <http://www.analyzedirect.com/>). The majority of freely available medical research software approaches have been designed for analysis of the brain, such as SPM <http://www.fil.ion.ucl.ac.uk/spm/software/spm2/> and Internet Image Viewer <http://james.psych.umn.edu/iiV/> [31], and not cardiovascular applications.

In summary, Segment differs compared to previous approaches in that to our knowledge it is the first source code extendable software application dedicated for cardiac image analysis. It is also to our knowledge by far the largest solution for applied medical imaging written in Matlab, a computer language that is widely used for image processing research. The strength of Segment lies in the combination of both a clinically applicable tool and a tool that easily can be further expanded by image processing experts and directly be used for clinical research.

#### **Predicted use of Segment**

We predict that Segment will continue to be used by cardiovascular researchers and in research groups with engineering teams. New functionality required to answer clinical research questions will be implemented and made available to the research community. The scripting functionality will enhance and facilitate larger clinical studies that are required to help cardiovascular imaging, and cardiovascular MRI in particular, to become an outcome-based imaging modality in medicine. Our hope is that Segment will function as a bridge between researchers in the field of image processing and researchers in cardiovascular research, both clinical and pre-clinical.

#### **Importance of validation**

Accurate and careful validation is of crucial importance in cardiovascular research. One feature that makes this project stand out from many other freely available medical image analysis software packages is the careful scientific validation that is performed when developing the new algorithms used in the software. However, one must remember that a chain is no stronger than its weakest link, and this is certainly true of quantitative image analysis. One such example is that flow quantification by MRI is validated in Segment, but it has been shown that each MRI scanner is unique; different pulse sequences have to be individually validated for each scanner since some scanners can introduce large sources of errors [18]. Another example is automated segmentation algorithms, which need to be supervised by trained and experienced observers to achieve the highest accuracy in clinically relevant measures.

## Conclusions

Segment is a cardiovascular image analysis software package that has been used in over 40 peer-reviewed scientific publications, indicating that the software has had an impact on cardiovascular research. Segment is a well-validated and comprehensive package that is freely available in an open source format for research purposes.

## Availability and requirements

The project name is Segment and the project home page is <http://segment.heiberg.se>. Pre-compiled versions of the software will be made available for Windows and Linux. The Matlab source code version of the program requires Matlab R2008a or later. The software is known to run under Mac OS X, but at the moment this is not supported. Segment is freely available for academic investigational research use, provided that relevant original research publications related to the software are cited. The software is also free for educational purposes. The terms of the licence do not generally include trials paid by pharmaceutical companies. For commercial use, Segment is sold and supported by Medviso AB, Lund, Sweden. Individuals or organisations are not allowed to compile software products derived from Segment that are to be sold commercially or shipped together with other commercial products without the express written permission of Medviso AB.

## Abbreviations

DE-MRI: Delayed Enhanced Magnetic Resonance Imaging; MRI: Magnetic Resonance Imaging; LV: Left Ventricle; LVM: Left Ventricle Mass; PACS: Picture Archiving Communication System; SPECT: Single Photon Emission Computed Tomography; CT: Computed Tomography; PET: Positron Emission Tomography.

## Acknowledgements

The authors thank all the users of Segment who have contributed with bug reports and suggestions for future enhancements. We also thank all the individuals who have contributed to Segment with code: Erik Bergvall (mathematical routines for strain analysis), Andreas Sigfridsson (parts of multiplanar reformatting), Helen Soneson (image fusion, strain analysis and LV SPECT segmentation), Erik Södervall (debugging and improvements), Johannes Töger (debugging of flow coordinate systems and implementation of 3D flow tools) and Johan Ugander (debugging and implementation of the point curve tool). We are grateful to the Swedish Research Council, the Swedish Heart and Lung Foundation and Scania Region for financial support.

## Trademarks

- DICOM is a registered trademark of the National Electrical Manufacturers Association for its standards publications relating to digital communication of medical information.
- Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.
- Mac OS X is a registered trademark of Apple Inc.
- Matlab is a registered trademark of The MathWorks Inc.
- Windows is a registered trademark of Microsoft Corporation Inc.

## Authors' contributions

EH designed and programmed most of the software and wrote major parts of the paper. JS programmed large parts of the general object segmentation module, designed and implemented the test script, and revised the manuscript for important intellectual content. MU, MC, HE and HA conceived many of the functions in the software and revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests, with the following two exceptions. Einar Heiberg Ph.D. developed most of the Segment software described in this study. Segment is freely available for research use, and it is sold for commercial use by Medviso AB, Lund, Sweden, a company of which Dr. Heiberg is the founder and the major shareholder. Jane Sjögren is employed by Medviso AB on a part-time basis.

Received: 21 September 2009

Accepted: 11 January 2010 Published: 11 January 2010

## References

1. Wolf I, Vetter M, Wegner I, Bottger T, Nolden M, Schobinger M, Hastenteufel M, Kunert T, Meinzer HP: **The medical imaging interaction toolkit.** *Medical image analysis* 2005, **9**:594-604.
2. Engblom H, Hedström E, Heiberg E, Wagner GS, Pahlm O, Arheden H: **Rapid initial reduction of hyperenhanced myocardium after reperfused first myocardial infarction suggest recovery of the peri-infarction zone: One year follow-up by MRI.** *Circulation Cardiovascular Imaging* 2009, **4**:47-55.
3. Heiberg E, Engblom H, Engvall J, Hedström E, Ugander M, Arheden H: **Semi-automatic quantification of myocardial infarction from delayed contrast enhanced magnetic resonance imaging.** *Scand Cardiovasc J* 2005, **39**:267-275.
4. Heiberg E, Ugander M, Engblom H, Götberg M, Olivecrona GK, Erlinge D, Arheden H: **Automated quantification of myocardial infarction from MR images by accounting for partial volume effects: animal, phantom, and human study.** *Radiology* 2008, **246**:581-588.
5. Heiberg E, Wigström L, Carlsson M, Bolger AF, Karlsson M: **Time Resolved Three-dimensional Automated Segmentation of the Left Ventricle.** *IEEE Computers in Cardiology* 2005; Lyon, France 2005, 599-602.
6. Carlsson M, Osman NF, Ursell PC, Martin AJ, Saeed M: **Quantitative MR measurements of regional and global left ventricular function and strain after intramyocardial transfer of VM202 into infarcted swine myocardium.** *American journal of physiology* 2008, **295**:H522-532.
7. Carlsson M, Wilson M, Martin AJ, Saeed M: **Myocardial microinfarction after coronary microembolization in swine: MR imaging characterization.** *Radiology* 2009, **250**:703-713.
8. Furtado AD, Carlsson M, Wintermark M, Ordozav K, Saeed M: **Identification of residual ischemia, infarction, and microvascular impairment in revascularized myocardial infarction using 64-slice MDCT.** *Contrast media & molecular imaging* 2008, **3**:198-206.
9. Carlsson M, Ugander M, Heiberg E, Arheden H: **The quantitative relationship between longitudinal and radial function in left, right, and total heart pumping in humans.** *American journal of physiology* 2007, **293**: H636-644.
10. Carlsson M, Ubachs JF, Hedström E, Heiberg E, Jovinge S, Arheden H: **Myocardium at risk after acute infarction in humans on cardiac magnetic resonance: quantitative assessment during follow-up and validation with single-photon emission computed tomography.** *JACC Cardiovasc Imaging* 2009, **2**:569-576.
11. Engblom H, Hedström E, Heiberg E, Wagner GS, Pahlm O, Arheden H: **Size and transmural extent of first-time reperfused myocardial infarction assessed by cardiac magnetic resonance can be estimated by 12-lead electrocardiogram.** *Am Heart J* 2005, **150**:920.
12. Carlsson M, Ursell PC, Saloner D, Saeed M: **Multidetector computed tomography for characterization of calcium deposits in reperfused myocardial infarction.** *Acta Radiol* 2009, **50**:396-405.
13. Carlsson M, Martin AJ, Ursell PC, Saloner D, Saeed M: **Magnetic resonance imaging quantification of left ventricular dysfunction following coronary microembolization.** *Magn Reson Med* 2009, **61**:595-602.

14. Hedstrom E, Engblom H, Frogner F, Astrom-Olsson K, Ohlin H, Jovinge S, Arheden H: **Infarct evolution in man studied in patients with first-time coronary occlusion in comparison to different species - implications for assessment of myocardial salvage.** *J Cardiovasc Magn Reson* 2009, **11**:38.
15. Wu EX, Wu Y, Nicholls JM, Wang J, Liao S, Zhu S, Lau CP, Tse HF: **MR diffusion tensor imaging study of postinfarct myocardium structural remodeling in a porcine model.** *Magn Reson Med* 2007, **58**:687-695.
16. Gilson WD, Kraitman DL: **Cardiac magnetic resonance imaging in small rodents using clinical 1.5 T and 3.0 T scanners.** *Methods (San Diego, Calif)* 2007, **43**:35-45.
17. Heiberg E: **Automated Feature Detection in Multidimensional Images.** Thesis Linköping University, Department of Biomedical Engineering 2004http://www.heiberg.se/einar/publications/pdf\_open/phd.pdf.
18. Chernobelsky A, Shubayev O, Comeau CR, Wolff SD: **Baseline correction of phase contrast images improves quantification of blood flow in the great vessels.** *J Cardiovasc Magn Reson* 2007, **9**:681-685.
19. Bernstein M, Zhou X, Polzin J, King K, Ganin A, Pelc N, Glover G: **Concomitant gradient terms in phase contrast MR: analysis and correction.** *Magnetic Resonance in Medicine* 1998, **39**:300-308.
20. Petzina R, Ugander M, Gustafsson L, Engblom H, Sjogren J, Hetzer R, Ingemansson R, Arheden H, Malmjö M: **Hemodynamic effects of vacuum-assisted closure therapy in cardiac surgery: assessment using magnetic resonance imaging.** *The Journal of thoracic and cardiovascular surgery* 2007, **133**:1154-1162.
21. Arheden H, Saeed M, Higgins CB, Gao DW, Bremerich J, Wyttenbach R, Dae MW, Wendland MF: **Measurement of the distribution volume of gadopentetate dimeglumine at echo-planar MR imaging to quantify myocardial infarction: comparison with 99mTc-DTPA autoradiography in rats.** *Radiology* 1999, **211**:698-708.
22. Engblom H, Carlsson MB, Hedstrom E, Heiberg E, Ugander M, Wagner GS, Arheden H: **The endocardial extent of reperfused first-time myocardial infarction is more predictive of pathologic Q waves than is infarct transmural: a magnetic resonance imaging study.** *Clin Physiol Funct Imaging* 2007, **27**:101-108.
23. Heiberg E, Engblom H, Ugander M, Arheden H: **Automated Calculation of Infarct Transmurality.** *IEEE Computers in Cardiology; Durham, USA* 2007, **165**-168.
24. Nilsson B, Heyden A: **A fast algorithm for level set-like active contours.** *Pattern Recognition Letters* 2003, **24**:1331-1337.
25. Svensson J, Gårdhagen R, Heiberg E, Ebbens T, Loyd D, Länne T, Karlsson M: **Feasibility of Patient Specific Aortic Blood Flow CFD Simulation.** *MICCAI; Copenhagen* 2006, **77**-86.
26. Sonesson H, Ubachs JF, Ugander M, Arheden H, Heiberg E: **An Improved Method for Automatic Segmentation of the Left Ventricle in Myocardial Perfusion SPECT.** *J Nucl Med* 2009, **50**:205-213.
27. Ugander M, Sonesson H, Heiberg E, Engblom H, Pals Jvd, Erlinge D, Arheden H: **A novel method for quantifying myocardial perfusion SPECT defect size by co-registration and fusion with MRI - an experimental ex vivo imaging pig heart study.** *Scand Cardiovasc J, Malmö* 2008, **42**(Suppl):47.
28. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, Pennell DJ, Rumberger JA, Ryan T, Verani MS: **Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association.** *Circulation* 2002, **105**:539-542.
29. Cain PA, Ugander M, Palmer J, Carlsson M, Heiberg E, Arheden H: **Quantitative polar representation of left ventricular myocardial perfusion, function and viability using SPECT and cardiac magnetic resonance: initial results.** *Clin Physiol Funct Imaging* 2005, **25**:215-222.
30. Gering DT, Nabavi A, Kikinis R, Hata N, O'Donnell LJ, Grimson WE, Jolesz FA, Black PM, Wells WM: **An integrated visualization system for surgical planning and guidance using image fusion and an open MR.** *J Magn Reson Imaging* 2001, **13**:967-975.
31. Lee JT, Munch KR, Carls JV, Pardo JV: **Internet image viewer (iiv).** *BMC medical imaging* 2008, **8**:10.

# Pre-publication history

The pre-publication history for this paper can be accessed here: <http://www.biomedcentral.com/1471-2342/10/1/prepub>

doi:10.1186/1471-2342-10-1

**Cite this article as:** Heiberg et al.: Design and validation of Segment - freely available software for cardiovascular image analysis. *BMC Medical Imaging* 2010 **10**:1.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)





## Paper II





## Research Article

# Validation and Development of a New Automatic Algorithm for Time-Resolved Segmentation of the Left Ventricle in Magnetic Resonance Imaging

Jane Tufvesson,<sup>1,2</sup> Erik Hedström,<sup>1,3</sup> Katarina Steding-Ehrenborg,<sup>1</sup> Marcus Carlsson,<sup>1</sup> Håkan Arheden,<sup>1</sup> and Einar Heiberg<sup>1,2,4</sup>

<sup>1</sup>Department of Clinical Physiology, Lund University Hospital, Lund University, 221 85 Lund, Sweden

<sup>2</sup>Department of Numerical Analysis, Centre for Mathematical Sciences, Faculty of Engineering, Lund University, 221 00 Lund, Sweden

<sup>3</sup>Department of Diagnostic Radiology, Lund University Hospital, Lund University, 221 85 Lund, Sweden

<sup>4</sup>Department of Biomedical Engineering, Faculty of Engineering, Lund University, 221 00 Lund, Sweden

Correspondence should be addressed to Einar Heiberg; [enar.heiberg@med.lu.se](mailto:enar.heiberg@med.lu.se)

Received 1 August 2014; Accepted 12 January 2015

Academic Editor: Peter M. A. Van Ooijen

Copyright © 2015 Jane Tufvesson et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Introduction.** Manual delineation of the left ventricle is clinical standard for quantification of cardiovascular magnetic resonance images despite being time consuming and observer dependent. Previous automatic methods generally do not account for one major contributor to stroke volume, the long-axis motion. Therefore, the aim of this study was to develop and validate an automatic algorithm for time-resolved segmentation covering the whole left ventricle, including basal slices affected by long-axis motion. **Methods.** Ninety subjects imaged with a cine balanced steady state free precession sequence were included in the study (training set  $n = 40$ , test set  $n = 50$ ). Manual delineation was reference standard and second observer analysis was performed in a subset ( $n = 25$ ). The automatic algorithm uses deformable model with expectation-maximization, followed by automatic removal of papillary muscles and detection of the outflow tract. **Results.** The mean differences between automatic segmentation and manual delineation were EDV  $-11$  mL, ESV  $1$  mL, EF  $-3\%$ , and LVM  $4$  g in the test set. **Conclusions.** The automatic LV segmentation algorithm reached accuracy comparable to interobserver for manual delineation, thereby bringing automatic segmentation one step closer to clinical routine. The algorithm and all images with manual delineations are available for benchmarking.

## 1. Introduction

Cardiovascular magnetic resonance (CMR) imaging can provide diagnostic information about the left ventricle (LV) with clinical parameters such as end-diastolic volume (EDV), end-systolic volume (ESV), ejection fraction (EF), left ventricular mass (LVM), stroke volume (SV), cardiac output (CO), peak ejection rate, peak filling rate, and regional wall thickening. To extract these clinical parameters current clinical practice is to perform endocardial and epicardial delineations manually, which is time consuming and therefore often only performed in end-diastole and end-systole [1]. However, delineations in two frames only will not give peak filling rate and peak ejection rate which require time-resolved segmentation. There is

also a need for segmentation throughout the cardiac cycle in the evaluation of patients with dyssynchrony, for example, to determine first and last segments with contraction [2]. With a typical time resolution of 30 frames per heartbeat, time-resolved manual delineation thus requires 15 times longer than manual delineation in only end-diastole and endsystole.

Automatic segmentation is desirable to reduce both analysis time and observer dependency. The continued need for manual delineation indicates that previously suggested automatic methods do not give satisfactory results. Often they do not cover the whole LV and there is a need for much manual interaction. Petitjean and Dacher [3] pointed out that it is hard to conclude on superiority of any of the previously proposed methods since the results are obtained

on images with different quality and in different patient populations. Also, the methods are validated using different error measurements, both clinical parameters and image processing error measurements. In midventricular slices the errors were by Petitjean and Dacher concluded to be generally satisfactory [3]. However, basal and apical slices generally yield higher errors [4].

Inclusion of all basal slices in the segmentation is important since the atrioventricular plane displacement is a major contributor to cardiac pumping [5, 6]. The long-axis motion causes the outflow tract to move in and out of the imaging plane during a cardiac cycle. Thereby, segmentation of endocardial and epicardial borders become more difficult in the most basal slices and an automatic detection of the long-axis motion is needed. To our knowledge three studies have included slices with outflow tract [4, 7, 8]. However, in the study by Jolly et al. [4] the detection of outflow tract was not defined, in the study by Hu et al. [7] the outflow tract was detected but the detection of long-axis motion was not defined, and finally in the study by Codella et al. [8] the user defined the most basal slice in both end-diastole and end-systole and thus the long-axis motion was not detected by the algorithm.

The aims of this study were (1) to develop an algorithm for time-resolved LV segmentation covering the whole LV, from the basal slices with outflow tract to the apex, and (2) to validate this new algorithm with regard to clinical parameters and image processing errors for comparison to previous algorithms, and (3) to provide software as well as images with manual delineation to enable benchmarking for future algorithms.

## 2. Methods

**2.1. Study Population and Design.** In total 90 subjects were included in the study, both patients referred for clinical evaluation of known or suspected coronary artery disease as well as healthy subjects and athletes. The subjects were scanned using a 1.5T MR scanner (Philips Intera CV, Philips, Best, The Netherlands) with a cardiac synergy coil. The sequence used was a balanced steady state free precession (bSSFP) sequence with retrospective ECG triggering. Typical imaging parameters were repetition time 2.8 ms, echo time 1.4 ms, flip angle 60°, SENSE factor of 2, spatial reconstructed resolution of  $1.4 \times 1.4 \times 8$  mm, and 30 reconstructed time frames per cardiac cycle (acquired spatial resolution  $2.3 \times 2.7 \times 8$  mm and temporal resolution 50 ms).

The subjects were divided into a training set ( $n = 40$ ) and a test set ( $n = 50$ ). The training set was used for the development and optimization of the algorithm, and the test set was used to validate the algorithm. The training set consists of 20 patients, 13 healthy volunteers, and 7 athletes. The test set consists of 20 patients, 20 healthy volunteers, and 10 athletes. Manual segmentation was performed for all slices in end-diastole and end-systole in both the training set and the test set by an experienced clinician (ErH with 14 years of CMR experience). The training set was reviewed for consensus by another experienced clinician (HA with 20 years of CMR experience). A subset of 25 subjects from the

test set (10 patients, 10 healthy volunteers, and 5 athletes) was used for second observer analysis, by another experienced clinician (MC with 14 years of CMR experience).

Image quality was representative of images likely found in daily clinical routine. Differences in clinical left ventricular parameters EDV, ESV, EF, LVM, SV, and CO between patients and healthy subjects in the test set were nonsignificant for all parameters except SV. The training set and test set with manual delineations are available upon request to the corresponding author to enable direct comparison to other methods.

**2.2. Automatic Segmentation Algorithm.** An automatic algorithm was developed for time-resolved segmentation of the endocardial and epicardial borders of the LV covering all ventricular slices from the most basal slices with outflow tract to the apex. The user input required by the algorithm is the definition of slices to analyze as the most basal slice and most apical slice containing any myocardium. The slices to analyze were in this study automatically selected to be the same as selected for the reference manual delineation. The algorithm was implemented in the freely available cardiac image analysis software Segment (<http://segment.heiberg.se/>) [9].

The algorithm is based on a deformable model framework. Deformable model is a segmentation method based on the idea of deforming a model to the location and shape of minimal energy in a force field. The model to deform is in this study a model of either the endocardial or the epicardial border. The initialization of the model is based on the image to segment and the initialization is further described in Step 3 of the algorithm. The force field which deforms the model consists of a weighted sum of image-dependent and model-dependent forces. The image-dependent forces are a balloon force image, calculated from image intensities and an edge force image, based on edge detection. The model-dependent forces are based on the curvature within the slice, between adjacent slices and between time frames of the cardiac cycle. The weighting of the forces in the deformable model was optimized based on the training set to obtain parameters suitable for the image type and quality in the training set. The optimization is further described in Deformable Model Optimization section.

Step 1 of the automatic segmentation algorithm is to define the center of the left ventricle, which is needed to calculate the balloon images in Step 2 and to initialize the deformable model in Step 3. Steps 4 and 5 use the deformable model for endocardial and epicardial segmentation. In Steps 6–8 the segmentation resulting from the deformable model is modified to account for the papillaries and the outflow tract. All steps are further described below.

The steps of the algorithm are as follows:

- (1) definition of the left ventricular center point,
- (2) calculation of balloon image,
- (3) initialization of segmentation,
- (4) endocardial segmentation,
- (5) epicardial segmentation,

- (6) exclusion of detached papillaries,
- (7) detection of outflow tract,
- (8) exclusion of attached papillaries.

*Definition of the Left Ventricular Center Point (Step 1).* First the center of the whole heart is defined from the largest bright region by smoothing and thresholding the image. The center of the left ventricular cavity is then defined as the center of gravity of the large, bright region closest to the right of the whole heart center point.

*Calculation of Balloon Image (Step 2).* The balloon force, which is the most important part of the deformable model, is defined using an expectation maximization (EM) algorithm. The balloon force drives the expansion and contraction of the curve and thereby should be a distinction between what to include and exclude in the endocardial and epicardial segmentations. The balloon image is mapped from the image intensities by estimating the distribution of intensities in the images. For endocardial segmentation the intensity distributions for blood and myocardium are estimated. In addition for epicardial segmentation, the intensity distribution of tissues surrounding the left ventricle is estimated. An EM-algorithm was utilized to estimate assumed Gaussian distribution of intensities for blood, myocardium, and surrounding tissues. As an initialization to the EM-algorithm, the mean and standard deviation for the intensity of blood were estimated in a cylinder with radius of 10 mm placed at the left ventricular center point. The endocardial balloon image was calculated as the Gaussian distribution for blood divided by the sum of the Gaussian distributions for blood and myocardium. The epicardial balloon image was calculated as the Gaussian distribution for myocardium divided by the sum of the Gaussian distributions for blood, myocardium, and surrounding tissues. The balloon force is positive for intensity values to include and negative for intensities to exclude and the balloon force was rescaled to the interval  $-1$  to  $1$ . Figure 1 shows the results from calculation of the balloon image.

*Initialization of Segmentation (Step 3).* To initialize both the endocardial and epicardial segmentations the endocardial balloon image is used. The endocardium is initialized at an estimated midmural center line and the epicardium is initialized as an estimate of the epicardial border. The initialization is divided into five substeps.

- (1) Thresholding the endocardial balloon image at zero to find regions representative of blood. Balloon force zero is representative of the probability of myocardium being equal to the probability of blood.
- (2) Finding the left ventricular blood pool as a region in the thresholded image which surrounds the left ventricular center point.
- (3) Estimating the endocardial border as the convex hull of the left ventricular blood pool. The convex hull is an estimation of the endocardial border excluding papillaries.

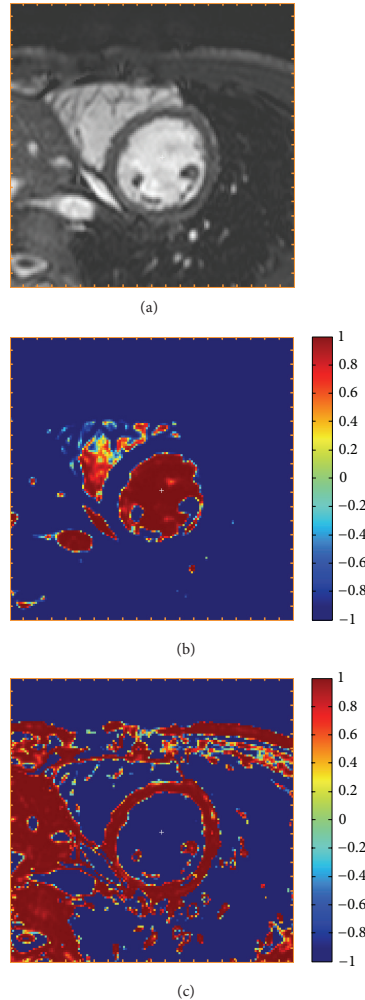


FIGURE 1: Calculation of balloon force (Step 2). A midventricular slice of a short-axis stack (a) and the endocardial (b) and epicardial (c) balloon force images calculated in Step 2 with the automatic algorithm. The color scale indicates how the deformable model should expand to include pixels with positive values (red) and contract to exclude pixels with negative values (blue).

- (4) Estimating the left ventricular wall thickness in each time frame by finding the mean distance from the initial curve to the right ventricular blood pool.
- (5) (a) Expanding the endocardial border estimated in Step 3 by a half wall thickness to get the endocardial initialization.

(b) Expanding the endocardial border estimated in Step 3 by one full wall thickness to get the epicardial initialization.

Figure 2 shows the initialization of endocardium and epicardium.

*Endocardial Segmentation (Step 4).* For endocardial segmentation, the deformable model is used with endocardial initialization, endocardial balloon force, and weighting of the forces optimized for endocardial segmentation. The deformable model formalism used has previously been described [10]. In short, in the deformable model, the node forces are normalized and projected onto the curve normal and the parameterization of the node points is kept equidistant. The deformable model includes balloon force, edge force, curvature force, temporal acceleration, and damping forces.

*Epicardial Segmentation (Step 5).* For epicardial segmentation, the deformable model is used with epicardial initialization, epicardial balloon force, and weighting of the forces optimized for epicardial segmentation. The epicardial balloon force is negative for blood and other tissues surrounding the myocardium and hence the deformable model will contract to not include any blood. To get an epicardial segmentation which expands outwards from the endocardial segmentation the epicardial balloon force was modified to be zero for all pixels inside the endocardium.

*Exclusion of Detached Papillaries (Step 6).* For measurement of ventricular volumes, the clinical standard is to exclude the papillaries from the myocardium and therefore the algorithm should also exclude the papillaries. Since papillary muscles have the same intensity as myocardium and the main driving force in the deformable model, the balloon force, is based on intensity, the algorithm may have difficulties with excluding the papillaries from the myocardium hence including the papillaries within the endocardial segmentation. The exclusion of papillaries is divided into two steps, this step and Step 8. In this step, detached papillaries are included inside the endocardial segmentation by taking the convex hull of the endocardial segmentation and refining the segmentation. The segmentation is refined by using the deformable model with a modified endocardial balloon force. The endocardial balloon force is modified by setting the balloon force to zero for papillaries, which are detected as pixels inside the convex hull with a negative balloon force.

*Detection of Outflow Tract (Step 7).* The deformable model gives endocardial and epicardial segmentation in all selected slices and time frames. Thereafter, long-axis motion and outflow tract are detected and in the basal slices the segmentation is adjusted accordingly. The detection of the long-axis motion is based on detecting sectors in the basal slices for which the intensities between the endocardial and epicardial segmentation are not typical for myocardium and sectors with a mean wall thickness of less than 2 millimeters. Basal slices were for detection of outflow tract defined as the most basal 40% of the ventricular length in end-diastole and all slices were divided into 24 sectors circumferentially.

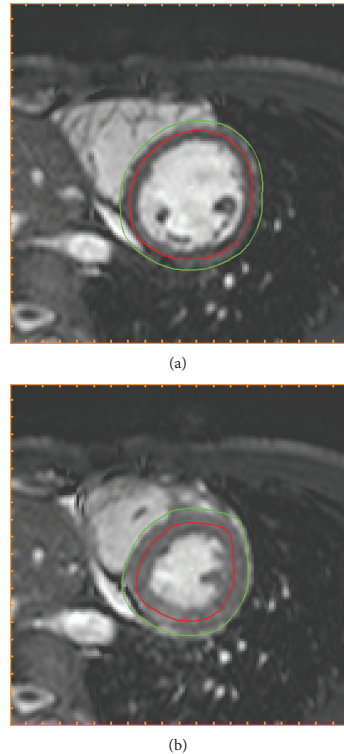


FIGURE 2: Initialization of segmentation (Step 3). The initializations of endocardial (red) and epicardial (green) borders resulting from Step 3 in the algorithm, shown in end-diastole (a) and end-systole (b) in the midventricular slice also used for Figure 1. The endocardial initialization is an estimation of the midmural line and the epicardial initialization is an estimation of the epicardial border.

The intensities in basal slices are compared to intensities in all slices. Sectors with a mean intensity 2 SD above the mean are marked as sectors to remove. Sectors can only be marked as sectors to remove if the sectors are also removed in a more basal slice. Sectors to be removed are smoothed over time and circumferentially in each slice and a morphological opening is performed to get a cohesive region to remove. To remove the marked sectors a straight line is drawn for both endocardium and epicardium. Thereby a D-shaped segmentation is obtained after adjustment for presence of outflow tract. Figure 3 shows the segmentation in a basal slice before and after the detection of outflow tract.

*Exclusion of Attached Papillaries (Step 8).* To exclude papillaries which are closely attached to the left ventricular wall in the segmentation, it is not sufficient to take convex hull and refine as in Step 6 since there is no blood volume which

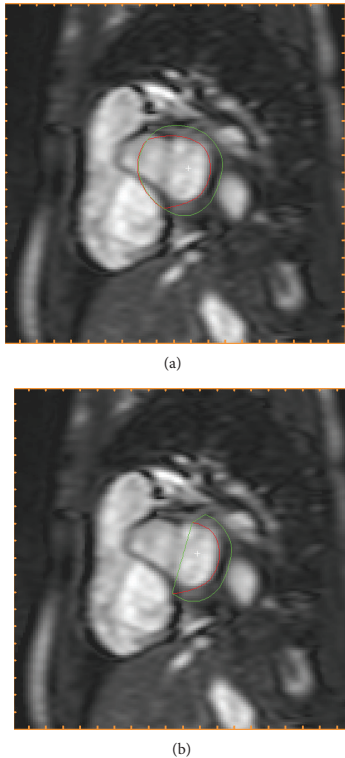


FIGURE 3: Detection of outflow tract (Step 7). The endocardial (red) and epicardial (green) segmentations are shown prior to the detection of outflow tract in Step 7 (a) and after adjustment of segmentation for presence of outflow tract (b) in the same basal slice in end-diastole.

can guide the deformable model on where to expand the segmentation. Therefore, in this step an expansion of the endocardial segmentation is calculated based on a constant papillary volume over time and a similar position of the papillary muscles over time. Sectors with a lower papillary volume inside the endocardial segmentation than in end-diastole are expanded to include more papillary volume. The long-axis displacement found when detecting the outflow tract in Step 7 is used to map slices in end-diastole to the corresponding slice in all other timeframes. Expansion of the endocardial segmentation is restricted to slices below the outflow tract in order to not falsely take the mitral valve into account as papillary muscle.

**2.3. Deformable Model Optimization.** Weighting of the forces in the deformable model was optimized with a steepest-descent method in a 2-factorial design by using the images in the training set with manual delineation as reference

standard. For the endocardial segmentation the error to minimize was the sum of the relative errors of the end-diastolic volume, and the relative number of falsely segmented pixels in end-diastole by comparing the automatic segmentation to manual delineation. Only the end-diastolic errors were included since the errors in end-systole are largely influenced by the presence of papillary muscles which is not especially accounted for within the deformable model.

For the epicardial segmentation the error to minimize in the optimization was the sum of the relative errors of left ventricular mass, in end-diastole and end-systole, and the relative number of falsely segmented pixels, in end-diastole and end-systole. In order to not take into account any volumetric errors in left ventricular mass given by the automatic endocardial segmentation, the left ventricular mass was during optimization calculated using the manual delineation of endocardium.

**2.4. Statistical Analyses.** In the test set the difference between manual delineation and automatic segmentation was computed for the clinical parameters EDV, ESV, EF, LVM, SV, and CO as well as the image processing error measurements dice similarity coefficient (DSC) [11] and point to curve distance (P2C).

The errors for clinical parameters are given both as absolute errors and as percentage of the result from the manual delineation. Paired *t*-test was performed with significance level  $P < 0.05$  to test for difference compared to manual delineations. A linear regression was performed for the clinical parameters and a regression *R*-value and corresponding *P*-value were calculated. The DSC is calculated as two times the volume of the intersection of two regions divided by the sum of the volume for those regions [11]. The DSC is therefore 0 if the regions do not overlap and 1 if the regions overlap perfectly. The P2C error was calculated as the distance between two borders in each slice and time frame where both borders were present. To calculate the distance both borders were resampled to be represented by 80 points spaced at every 4.5 degrees. The DSC and P2C errors were calculated between automatic segmentation and manual delineation for both endocardial and epicardial segmentation separately. The DSC and P2C error were calculated as a mean over all slices in both end-diastole and end-systole as well as separately for end-diastole and end-systole and separately for basal, midventricular, and apical slices. Basal, midventricular, and apical slices were defined as one third each of the ventricular length in both end-diastole and end-systole. All errors were reported as mean  $\pm$  SD.

In the subset for which second observer manual delineation was performed the same error calculations as for the full test set were performed for automatic segmentation versus reference manual delineation and for second observer manual delineation versus reference manual delineation.

### 3. Results

Automatic segmentation was performed and compared to manual delineation in the test and compared to interobserver variability in a second observer subset. In one patient



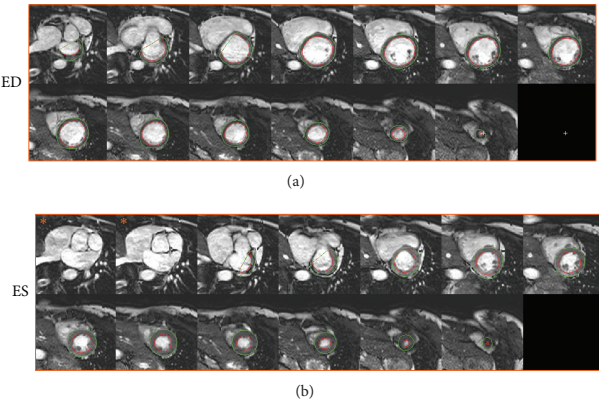


FIGURE 4: Example of segmentation in end-diastole and end-systole. An example of automatic segmentation is shown in end-diastole (a) and end-systole (b). Each panel shows the short axis stack covering the left ventricle from base to apex with endocardial (red) and epicardial (green) segmentations. Note how the outflow tract has moved out of the two most basal slices in end-systole (b, images marked \*), compared to end-diastole (a) and that the algorithm has automatically corrected for this long-axis motion.

the automatic segmentation failed due to a severe bright fold-in artifact connecting the right and left ventricle. This patient was excluded from further analysis resulting in a test set of 49 patients and a second observer subset of 24 patients. Figure 4 shows an example of automatic segmentation in all slices in end-diastole and end-systole. A comparison between automatic segmentation and manual delineation can be seen in Figure 5 for a basal, midventricular, and apical slice in end-diastole and end-systole. In the additional file a time-resolved 3D-rendering of left ventricle shows the long-axis motion of the epicardial surface resulting from the automatic segmentation algorithm. The differences between automatic segmentation and manual delineation for clinical parameters were EDV  $-11 \pm 11$  mL ( $R = 0.96$ ), ESV  $1 \pm 10$  mL ( $R = 0.95$ ), EF  $-3 \pm 4\%$  ( $R = 0.86$ ), LVM  $4 \pm 15$  g ( $R = 0.87$ ), SV  $-12 \pm 8$  mL ( $R = 0.92$ ), and CO  $-0.7 \pm 0.5$  L/min ( $R = 0.94$ ) (Table 1, Figures 6 and 7). The image processing error measurements were for endocardial segmentation DSC =  $0.91 \pm 0.03$  and P2C =  $2.1 \pm 0.5$  mm and for epicardial segmentation DSC =  $0.93 \pm 0.02$  and P2C =  $2.1 \pm 0.5$  mm as mean over all slices and both end-diastole and end-systole (Table 2). End-diastolic image processing error measurements performed better than end-systolic (Table 2). Midventricular slices performed better than basal and apical slices (Table 3).

In the subset for second observer analysis the differences between second observer manual delineation and reference manual delineation were EDV  $10 \pm 4$  mL, ESV  $5 \pm 5$  mL, EF  $0 \pm 2\%$ , LVM  $-7 \pm 9$  g, SV  $5 \pm 6$  mL, and CO  $0.3 \pm 0.4$  L/min compared to the differences between automatic segmentation and the reference manual delineation which were EDV  $-9 \pm 10$  mL, ESV  $3 \pm 8$ , EF  $-3 \pm 3\%$ , LVM  $2 \pm 16$  g, SV  $-12 \pm 8$  mL, and CO  $-0.7 \pm 0.4$  L/min (Table 4). The results for the image processing error measurements DSC and P2C for the second observer subset are given in Tables 5 and 6.

TABLE 1: Clinical parameters in test set. Results for clinical parameters in the full test set ( $n = 49$ ) as differences between automatic segmentation and manual delineation.

	Absolute difference	Relative difference	P value
EDV	$-11 \pm 11$ mL	$-6 \pm 6\%$	$<0.01$
ESV	$1 \pm 10$ mL	$1 \pm 13\%$	0.57
EF	$-3 \pm 4\%$	$-5 \pm 7\%$	$<0.01$
LVM	$4 \pm 15$ g	$4 \pm 14\%$	0.07
SV	$-12 \pm 8$ mL	$-11 \pm 8\%$	$<0.01$
CO	$-0.7 \pm 0.5$ L/min	$-11 \pm 8\%$	$<0.01$

Absolute and relative values are expressed as mean  $\pm$  SD. EDV = end-diastolic volume, ESV = end-systolic volume, EF = ejection fraction, and LVM = left ventricular mass.

TABLE 2: Image processing error measurement in test set. Image processing error measurements in the full test set ( $n = 49$ ) as dice similarity coefficient (DSC) and point to curve (P2C) between automatic segmentation and manual delineation.

	Dice similarity coefficient (DSC)	Point to curve (P2C)
Endocardium overall	$0.91 \pm 0.03$	$2.1 \pm 0.5$ mm
Endocardium ED	$0.93 \pm 0.03$	$1.9 \pm 0.6$ mm
Endocardium ES	$0.85 \pm 0.04$	$2.3 \pm 0.5$ mm
Epicardium overall	$0.93 \pm 0.02$	$2.1 \pm 0.5$ mm
Epicardium ED	$0.94 \pm 0.02$	$2.1 \pm 0.6$ mm
Epicardium ES	$0.91 \pm 0.03$	$2.2 \pm 0.7$ mm

Differences are expressed as mean  $\pm$  SD. For a perfect overlap between the regions DSC should be 1 and P2C should be 0. ED = end diastole, ES = end systole.

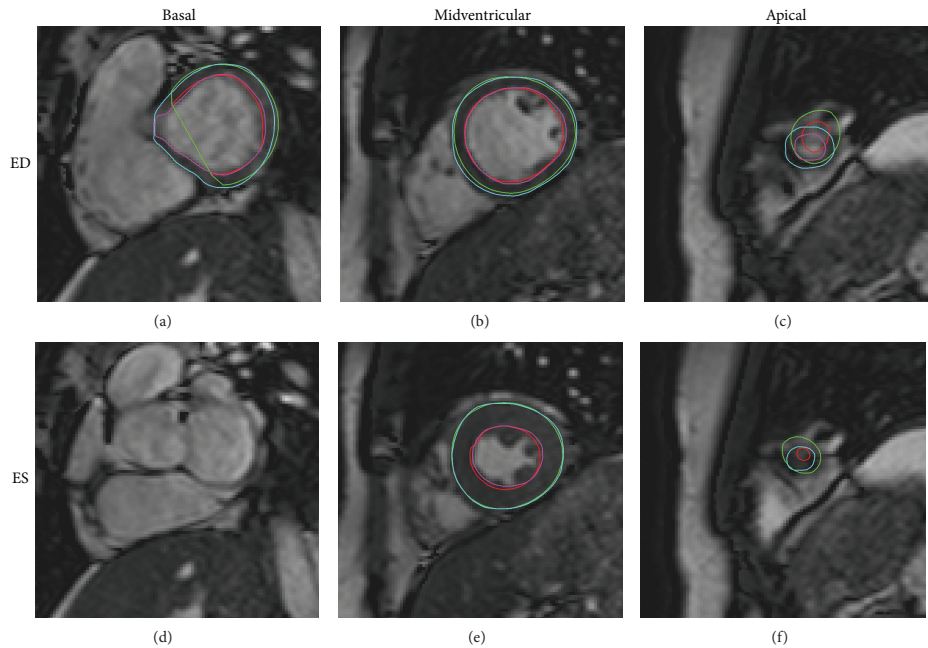


FIGURE 5: Automatic segmentation compared to manual delineation in a basal, midventricular and apical slice. Automatic segmentation (endocardium in red and epicardium in green) and manual delineation (endocardium in pink and epicardium in light blue) shown in end-diastole ((a), (b), and (c)) and end-systole ((d) (e), and (f)) for the most basal slice with outflow tract moving out of the imaging plane ((a), (d)), a midventricular slice with papillaries ((b), (e)) and an apical slice with minimal lumen in end-systole ((c), (f)).

TABLE 3: Image processing error measurements in test set divided into slice sections. Image processing error measurements in the full test set ( $n = 49$ ) as dice similarity coefficient (DSC) and point to curve (P2C) between automatic segmentation and manual delineation.

	Dice similarity coefficient (DSC)	Point to curve (P2C)
Endocardium basal	$0.88 \pm 0.06$	$2.7 \pm 1.0$ mm
Endocardium midventricular	$0.94 \pm 0.02$	$1.6 \pm 0.4$ mm
Endocardium apical	$0.89 \pm 0.03$	$2.1 \pm 0.7$ mm
Epicardium basal	$0.89 \pm 0.05$	$3.3 \pm 1.2$ mm
Epicardium midventricular	$0.96 \pm 0.02$	$1.3 \pm 0.5$ mm
Epicardium apical	$0.92 \pm 0.03$	$2.2 \pm 0.8$ mm

Differences are expressed as mean  $\pm$  SD. For a perfect overlap between the regions DSC should be 1 and P2C should be 0. Basal, midventricular, and apical sections are defined as 1/3 each of the ventricular length in end diastole and end systole separately.

4. Discussion

We have developed an automatic algorithm for time-resolved LV segmentation in magnetic resonance cine balanced steady

state free precession (MRSSFP) images. The segmentation is performed in all time frames and all ventricular slices including the slices in which the mitral valve plane and outflow tract move in and out of the slice during a heartbeat. The only manual user input is definition of the most basal and most apical slices including any myocardium in end-diastole. This study brings a state-of-the-art left ventricle segmentation tool to applied clinical research, as the software and source code are provided in open access to researchers. Furthermore, both algorithm and images with ground truth manual delineations are made available for benchmark against future LV segmentation algorithms.

The major algorithmic contributions towards a clinically applicable automatic segmentation method in this study is (1) the use of an EM-algorithm to calculate the distinction between blood, myocardium, and tissues surrounding the heart, (2) removal of papillary muscles by convex hull expansion and expansion to get constant papillary volume, (3) the detection of the outflow tract when moving in and out of the imaging plane, and (4) usage of an optimization step to tune otherwise arbitrary set parameters to the images used.

The algorithm was validated in a test set of 49 subjects and both the clinical parameters, EDV, ESV, EF, and LVM, and the image processing error measurements, DSC and P2C, were

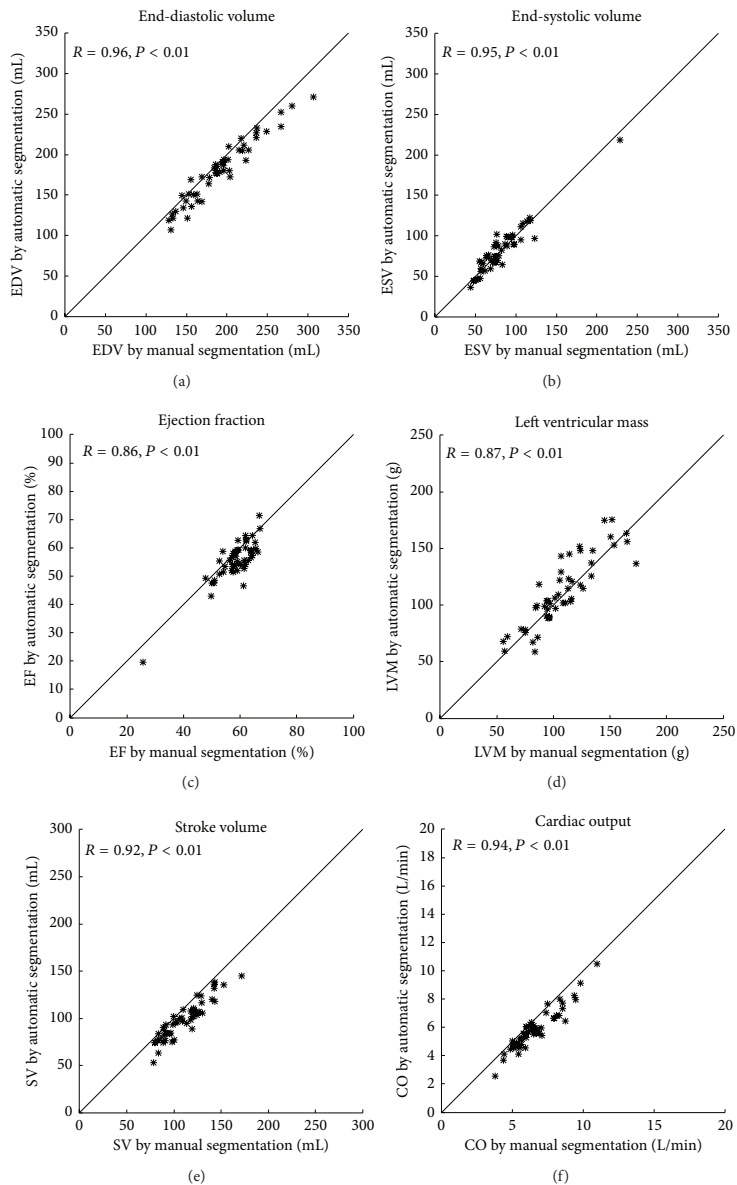


FIGURE 6: Correlations between automatic segmentation and manual delineation in the test set. Automatic segmentation plotted against manual delineation for end-diastolic volume (EDV, (a)), end-systolic volume (ESV, (b)), ejection fraction (EF, (c)), left ventricular mass (LVM, (d)), stroke volume (SV, (e)) and cardiac output (CO, (f)). The line indicates the line of identity.



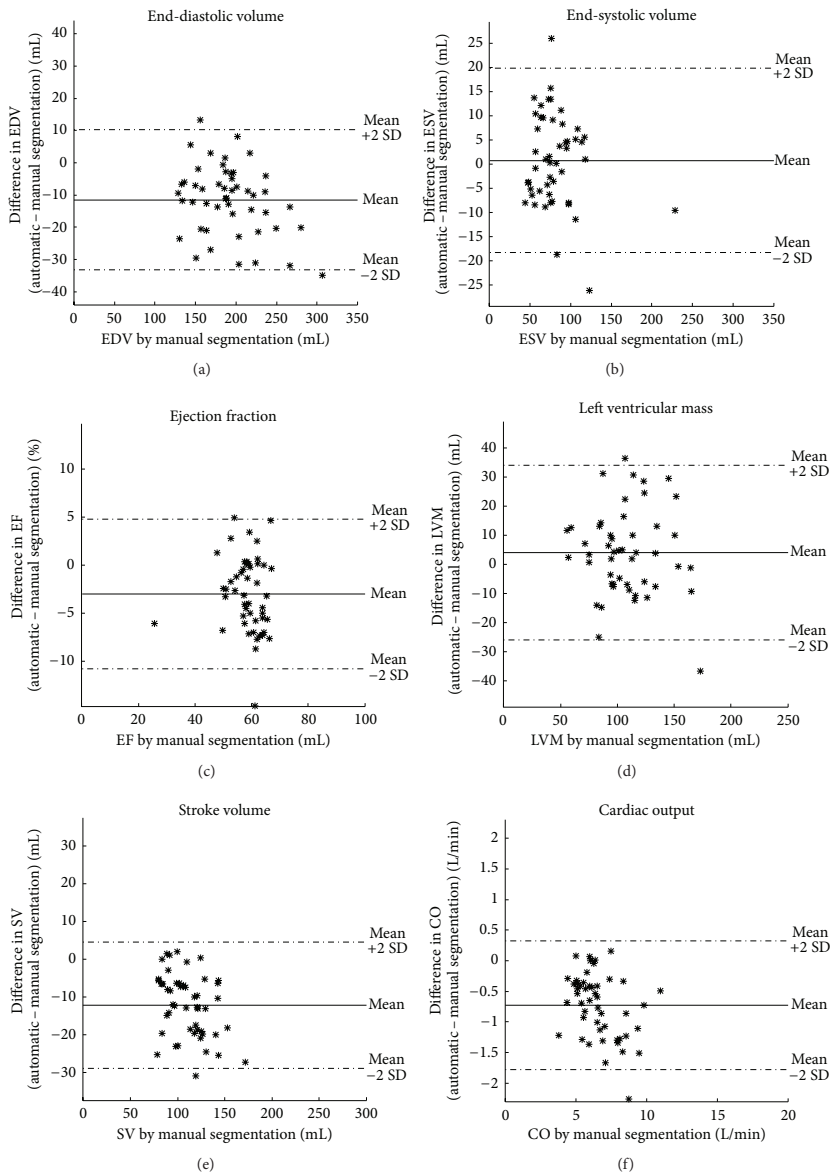


FIGURE 7: Bias between automatic segmentation and manual delineation in the test set. Differences between automatic segmentation and manual delineation plotted against manual delineation for end-diastolic volume (EDV, (a)), end-systolic volume (ESV, (b)), ejection fraction (EF, (c)), left ventricular mass (LVM, (d)), stroke volume (SV, (e)) and cardiac output (CO, (f)). Solid line indicates mean and dashed lines indicate mean  $\pm$  2SD.

TABLE 4: Clinical parameters in second observer subset. Differences for clinical parameters in the second observer subset ( $n = 24$ ) for second observer manual delineation versus manual reference delineation and for automatic segmentation versus manual reference delineation.

	Automatic segmentation versus manual reference			Second observer versus manual reference		
	Absolute difference	Relative difference	P value	Absolute difference	Relative difference	P value
EDV	$-9 \pm 10$ mL	$-5 \pm 5\%$	$<0.01$	$10 \pm 4$ mL	$6 \pm 2\%$	$<0.01$
ESV	$3 \pm 8$ mL	$4 \pm 12\%$	0.1	$5 \pm 5$ mL	$6 \pm 6\%$	$<0.01$
EF	$-3 \pm 3\%$	$-6 \pm 6\%$	$<0.01$	$0 \pm 2\%$	$-1 \pm 4\%$	0.44
LVM	$2 \pm 16$ g	$3 \pm 13\%$	0.55	$-7 \pm 9$ g	$-7 \pm 8\%$	$<0.01$
SV	$-12 \pm 8$ mL	$-10 \pm 6\%$	$<0.01$	$5 \pm 6$ mL	$5 \pm 5\%$	$<0.01$
CO	$-0.7 \pm 0.4$ L/min	$-10 \pm 6\%$	$<0.01$	$0.3 \pm 0.4$ L/min	$5 \pm 5\%$	$<0.01$

Absolute and relative difference expressed as mean  $\pm$  SD. EDV = end-diastolic volume, ESV = end-systolic volume, EF = ejection fraction, and LVM = left ventricular mass.

TABLE 5: Image processing error measurements in second observer subset. Image processing error measurements in the second observer subset ( $n = 24$ ) as dice similarity coefficient (DSC) and point to curve (P2C) for second observer manual delineation versus manual reference delineation and for automatic segmentation versus manual reference delineation.

	Automatic segmentation versus manual reference		Second observer versus manual reference	
	DSC	P2C	DSC	P2C
Endocardium overall	$0.91 \pm 0.02$	$2.0 \pm 0.4$ mm	$0.95 \pm 0.01$	$1.2 \pm 0.2$ mm
Endocardium ED	$0.93 \pm 0.02$	$1.8 \pm 0.5$ mm	$0.96 \pm 0.01$	$1.1 \pm 0.3$ mm
Endocardium ES	$0.85 \pm 0.04$	$2.4 \pm 0.5$ mm	$0.92 \pm 0.03$	$1.4 \pm 0.3$ mm
Epicardium overall	$0.93 \pm 0.01$	$2.2 \pm 0.4$ mm	$0.96 \pm 0.01$	$1.2 \pm 0.33$ mm
Epicardium ED	$0.94 \pm 0.01$	$2.0 \pm 0.5$ mm	$0.97 \pm 0.01$	$1.1 \pm 0.4$ mm
Epicardium ES	$0.91 \pm 0.02$	$2.4 \pm 0.6$ mm	$0.95 \pm 0.01$	$1.4 \pm 0.4$ mm

Difference are expressed as mean  $\pm$  SD. For a perfect overlap between the regions DSC should be 1 and P2C should be 0. ED = end diastole, ES = end systole.

TABLE 6: Image processing error measurements in second observer set divided into slice sections. Image processing error measurements in the second observer subset ( $n = 24$ ) as dice similarity coefficient (DSC) and point to curve (P2C) for second observer manual delineation versus manual reference delineation and for automatic segmentation versus manual reference delineation.

	Automatic segmentation versus manual reference		Second observer versus manual reference	
	DSC	P2C	DSC	P2C
Endocardium basal	$0.88 \pm 0.06$	$2.7 \pm 1.1$ mm	$0.94 \pm 0.02$	$1.5 \pm 0.4$ mm
Endocardium midventricular	$0.94 \pm 0.01$	$1.6 \pm 0.4$ mm	$0.96 \pm 0.01$	$1.1 \pm 0.3$ mm
Endocardium apical	$0.90 \pm 0.03$	$2.0 \pm 0.7$ mm	$0.94 \pm 0.01$	$1.1 \pm 0.3$ mm
Epicardium basal	$0.89 \pm 0.05$	$3.3 \pm 1.2$ mm	$0.95 \pm 0.02$	$1.5 \pm 0.6$ mm
Epicardium midventricular	$0.96 \pm 0.01$	$1.3 \pm 0.4$ mm	$0.97 \pm 0.01$	$0.8 \pm 0.2$ mm
Epicardium apical	$0.92 \pm 0.02$	$2.3 \pm 0.7$ mm	$0.95 \pm 0.01$	$1.4 \pm 0.5$ mm

Difference are expressed as mean  $\pm$  SD. For a perfect overlap between the regions DSC should be 1 and P2C should be 0. Basal, midventricular, and apical sections are defined as 1/3 each of the ventricular length in end diastole and end systole separately.

reported to allow comparison to errors reported in previous studies. The proposed algorithm has a DSC and P2C error similar to the ones reported in previous studies [3, 4, 12–14]. However, direct comparison between studies is difficult due to differences in methodology. In previous studies it is not defined either how the basal slices were selected, or if the basal slices were excluded or defined separately for end-diastole and end-systole thereby not including the long-axis motion. Furthermore, results may not be directly comparable due to differences in patient population and sequences used for imaging. For instance the test set in the MICCAI challenge [3] was acquired without parallel acquisition techniques which is now clinical standard. In the sequel challenge STACOM [14] not all results were derived using manual delineation as

ground truth. A new test set and training set were therefore acquired for this study in order to have images with parallel acquisition, covering all slices and with manual delineation as ground truth. In comparison to the present study, Codella et al. [8, 15] reported better results for all clinical parameters, which is expected with the higher level of user input used in their algorithm LV-METRIC. The present study has a low level of user input with only a selection of slices to include in segmentation. Hu et al. [7] developed a detection of the outflow tract and reported DSC and P2C similar to the present study. However, their method description does not define detection of the outflow tract moving out of the imaging plane. Since the long-axis motion is a major contributor to cardiac pumping [5] it is important to include

the basal slices and account for the contraction along the long-axis. Segmentation of the most basal slices with outflow tract becomes more difficult when the myocardium moves in and out of the imaging plane. The proposed algorithm includes all slices with results similar to previous studies not including all basal slices, which brings the algorithm one step closer to automatic LV segmentation applicable for the clinical routine.

The proposed algorithm was compared to interobserver variability of manual delineation in a subset as a major goal of automatic segmentation methods is to reduce observer dependency. The proposed method showed a bias comparable to interobserver variability by manual delineation for the clinical parameters, lower or similar bias for EDV, ESV, and LVM and higher bias for EF, SV, and CO. The SD for the clinical parameters was approximately twice the value found for interobserver variability. The interobserver variability measured as P2C error was overall 1.2 mm compared to 2 mm reported in a previous study [16]. The interobserver variability measured as clinical parameters was overall comparable to those reported in previous studies [17–19]. The standard deviation of LVM for interobserver variability was in this study 7 g which falls within the range of published values from 5 g in a normal material for gradient echo images [20] to 14 g in a study where bSSFP short axis delineations were compared to long-axis delineations [21]. The large range in interobserver variability measurements reported in the literature can most likely be explained by differences in methodology used in the basal regions, differences in image quality, and amount of consensus training. Again many of the studies report differences differently and direct comparisons are difficult.

In order for the algorithm to reach results fully comparable to interobserver variability between two experienced observers, further improvement is needed. By improving the use of the EM-algorithm and by improving the detection of papillary muscles and outflow tract both the accuracy and precision may be reduced. The algorithm might also be further improved to have a smoother segmentation over the cardiac cycle by using more than two time frames in the optimization of parameters and hence possibly get a higher weight on the time dependent parameter. As for all automatic segmentation algorithms a manual approval and possibly manual corrections are needed in a clinical setting.

A limitation to the study is that the training and test set used only patients with coronary artery disease. Other patient categories with, for example, left ventricular dyssynchrony or pronounced trabeculations may need special consideration in the algorithm and further validation.

## 5. Conclusion

We have developed an automatic algorithm for time-resolved segmentation of all LV slices containing any myocardium in magnetic resonance balanced steady state free precession images. The algorithm was quantitatively validated in 49 subjects and both algorithm and images with reference manual delineations are available for benchmark against future LV segmentation algorithms. The algorithm showed a bias comparable to interobserver variability between two

experienced observers for the clinical parameters EDV, ESV, EF, LVM, SV, and CO. With a dice and P2C error similar to previous studies the proposed algorithm is favorable due to low level of user input and automatic correction for long-axis motion. The algorithm is one step closer to an automatic segmentation applicable for clinical routine.

## Disclosure

Einar Heiberg is the founder of Medviso AB, Lund, Sweden, which sells a commercial version of the otherwise freely available software Segment. Jane Tufvesson is employed by Medviso AB on a part-time basis.

## Conflict of Interests

All other authors declare that there is no conflict of interests regarding the publication of this paper.

## Authors' Contribution

All authors contributed to concept and design of the study, input to the development of the automatic segmentation algorithm, and revised the paper. All authors read and approved the final paper. Jane Tufvesson developed and implemented the automatic segmentation algorithm, analyzed and interpreted segmentation results, and drafted the paper. Erik Hedström performed manual delineation in training set and test set. Katarina Steding-Ehrenborg included normal subjects and athletes to the study. Marcus Carlsson performed second observer manual delineation. Håkan Arheden performed consensus reading of the manual delineation in training set. Einar Heiberg conceived the study.

## Acknowledgments

The authors would like to acknowledge Ann-Helen Arvidsson and Christel Carlander, both with the Lund Cardiac MR Group, for their skillful assistance with image acquisition, and Joey Ubachs for inclusion of patients to the study. This study has been funded by the Swedish Research Council (2011-3916, 2008-2949, and 2012-4944), The Swedish Heart and Lung Foundation, The Medical Faculty of Lund University, Sweden, and Region of Scania, Sweden.

## References

- [1] J. Schulz-Menger, D. A. Bluemke, J. Bremerich et al., "Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) Board of Trustees Task Force on Standardized Post Processing," *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, article 35, 2013.
- [2] P. Foley, K. Khadjooi, J. Ward et al., "Radial dyssynchrony assessed by cardiovascular magnetic resonance in relation to left ventricular function, myocardial scarring and QRS duration in patients with heart failure," *Journal of Cardiovascular Magnetic Resonance*, vol. 11, no. 1, article 50, 2009.

- [3] C. Petitjean and J.-N. Dacher, "A review of segmentation methods in short axis cardiac MR images," *Medical Image Analysis*, vol. 15, no. 2, pp. 169–184, 2011.
- [4] M. P. Jolly, H. Xue, L. Grady, and J. Guehring, "Combining registration and minimum surfaces for the segmentation of the left ventricle in cardiac cine MR images," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 12, pp. 910–918, Springer, Berlin, Germany, 2009.
- [5] M. Carlsson, M. Ugander, H. Mosén, T. Buhre, and H. Arheden, "Atrioventricular plane displacement is the major contributor to left ventricular pumping in healthy adults, athletes, and patients with dilated cardiomyopathy," *The American Journal of Physiology—Heart and Circulatory Physiology*, vol. 292, no. 3, pp. H1452–H1459, 2007.
- [6] K. Steding-Ehrenborg, M. Carlsson, S. Stephensen, and H. Arheden, "Atrial aspiration from pulmonary and caval veins is caused by ventricular contraction and secures 70% of the total stroke volume independent of resting heart rate and heart size," *Clinical Physiology and Functional Imaging*, vol. 33, no. 3, pp. 233–240, 2013.
- [7] H. Hu, H. Liu, Z. Gao, and L. Huang, "Hybrid segmentation of left ventricle in cardiac MRI using gaussian-mixture model and region restricted dynamic programming," *Magnetic Resonance Imaging*, vol. 31, no. 4, pp. 575–584, 2013.
- [8] N. C. F. Codella, M. D. Cham, R. Wong et al., "Rapid and accurate left ventricular chamber quantification using a novel CMR segmentation algorithm: a clinical validation study," *Journal of Magnetic Resonance Imaging*, vol. 31, no. 4, pp. 845–853, 2010.
- [9] E. Heiberg, J. Sjögren, M. Ugander, M. Carlsson, H. Engblom, and H. Arheden, "Design and validation of segment—freely available software for cardiovascular image analysis," *BMC Medical Imaging*, vol. 10, article 1, 2010.
- [10] E. Heiberg, L. Wigstrom, M. Carlsson, A. F. Bolger, and M. Karlsson, "Time resolved three-dimensional automated segmentation of the left ventricle," in *Proceedings of the Computers in Cardiology*, pp. 599–602, Lyon, France, September 2005.
- [11] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [12] D. Grosgeorge, C. Petitjean, J. Caudron, J. Fares, and J. N. Dacher, "Automatic cardiac ventricle segmentation in MR images: a validation study," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 5, pp. 573–581, 2011.
- [13] Y. Huang, X. Zhou, B. Miao et al., "An image based system biology approach for Alzheimer's disease pathway analysis," in *Proceedings of the Life Science Systems and Applications Workshop (LiSSA '09)*, pp. 128–132, 2009.
- [14] A. Suinesiaputra, B. Cowan, J. P. Finn et al., "Left ventricular segmentation challenge from cardiac MRI: a collation study," in *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, O. Camara, E. Konukoglu, M. Pop, K. Rhode, M. Sermesant, and A. Young, Eds., vol. 7085 of *Lecture Notes in Computer Science*, pp. 88–97, Springer, Berlin, Germany, 2012.
- [15] N. C. F. Codella, H. Y. Lee, D. S. Fieno et al., "Improved left ventricular mass quantification with partial voxel interpolation in vivo and necropsy validation of a novel cardiac MRI segmentation algorithm," *Circulation: Cardiovascular Imaging*, vol. 5, no. 1, pp. 137–146, 2012.
- [16] H. C. van Assen, M. G. Danilouchkine, A. F. Frangi et al., "SPASM: a 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data," *Medical Image Analysis*, vol. 10, no. 2, pp. 286–303, 2006.
- [17] N. C. F. Codella, J. W. Weinsaft, M. D. Cham, M. Janik, M. R. Prince, and Y. Wang, "Left ventricle: automated segmentation by using myocardial effusion threshold reduction and intravoxel computation at MR imaging," *Radiology*, vol. 248, no. 3, pp. 1004–1012, 2008.
- [18] S. Plein, W. H. T. Smith, J. P. Ridgway et al., "Qualitative and quantitative analysis of regional left ventricular wall dynamics using real-time magnetic resonance imaging: comparison with conventional breath-hold gradient echo acquisition in volunteers and patients," *Journal of Magnetic Resonance Imaging*, vol. 14, no. 1, pp. 23–30, 2001.
- [19] R. J. van der Geest, B. P. F. Lelieveldt, E. Angelí et al., "Evaluation of a new method for automated detection of left ventricular boundaries in time series of magnetic resonance images using an active appearance motion model," *Journal of Cardiovascular Magnetic Resonance*, vol. 6, no. 3, pp. 609–617, 2004.
- [20] P. A. Cain, R. Ahl, E. Hedstrom et al., "Age and gender specific normal values of left ventricular mass, volume and function for gradient echo magnetic resonance imaging: a cross sectional study," *BMC Medical Imaging*, vol. 9, article 2, 2009.
- [21] H. Childs, L. Ma, M. Ma et al., "Comparison of long and short axis quantification of left ventricular volume parameters by cardiovascular magnetic resonance, with ex-vivo validation," *Journal of Cardiovascular Magnetic Resonance*, vol. 13, no. 1, article 40, 2011.

## Paper III



RESEARCH

Open Access

# Semi-automatic segmentation of myocardium at risk in T2-weighted cardiovascular magnetic resonance

Jane Sjögren<sup>1,2</sup>, Joey FA Ubachs<sup>1</sup>, Henrik Engblom<sup>1</sup>, Marcus Carlsson<sup>1</sup>, Håkan Arheden<sup>1</sup> and Einar Heiberg<sup>1\*</sup>

## Abstract

**Background:** T2-weighted cardiovascular magnetic resonance (CMR) has been shown to be a promising technique for determination of ischemic myocardium, referred to as myocardium at risk (MaR), after an acute coronary event. Quantification of MaR in T2-weighted CMR has been proposed to be performed by manual delineation or the threshold methods of two standard deviations from remote (2SD), full width half maximum intensity (FWHM) or Otsu. However, manual delineation is subjective and threshold methods have inherent limitations related to threshold definition and lack of *a priori* information about cardiac anatomy and physiology. Therefore, the aim of this study was to develop an automatic segmentation algorithm for quantification of MaR using anatomical *a priori* information.

**Methods:** Forty-seven patients with first-time acute ST-elevation myocardial infarction underwent T2-weighted CMR within 1 week after admission. Endocardial and epicardial borders of the left ventricle, as well as the hyper enhanced MaR regions were manually delineated by experienced observers and used as reference method. A new automatic segmentation algorithm, called Segment MaR, defines the MaR region as the continuous region most probable of being MaR, by estimating the intensities of normal myocardium and MaR with an expectation maximization algorithm and restricting the MaR region by an *a priori* model of the maximal extent for the user defined culprit artery. The segmentation by Segment MaR was compared against inter observer variability of manual delineation and the threshold methods of 2SD, FWHM and Otsu.

**Results:** MaR was  $32.9 \pm 10.9\%$  of left ventricular mass (LVM) when assessed by the reference observer and  $31.0 \pm 8.8\%$  of LVM assessed by Segment MaR. The bias and correlation was,  $-1.9 \pm 6.4\%$  of LVM,  $R = 0.81$  ( $p < 0.001$ ) for Segment MaR,  $-2.3 \pm 4.9\%$ ,  $R = 0.91$  ( $p < 0.001$ ) for inter observer variability of manual delineation,  $-7.7 \pm 11.4\%$ ,  $R = 0.38$  ( $p = 0.008$ ) for 2SD,  $-21.0 \pm 9.9\%$ ,  $R = 0.41$  ( $p = 0.004$ ) for FWHM, and  $5.3 \pm 9.6\%$ ,  $R = 0.47$  ( $p < 0.001$ ) for Otsu.

**Conclusions:** There is a good agreement between automatic Segment MaR and manually assessed MaR in T2-weighted CMR. Thus, the proposed algorithm seems to be a promising, objective method for standardized MaR quantification in T2-weighted CMR.

## Background

Myocardium at risk (MaR) is defined as the ischemic myocardium during coronary artery occlusion and is the region that will be subject to infarction if the blood flow is not restored. Myocardium at risk can be measured using T2-weighted cardiovascular magnetic resonance

(CMR) [1] due to the myocardial edema occurring in the ischemic myocardium [2,3] up to one week after percutaneous coronary intervention (PCI) [4]. By determining MaR using T2-weighted CMR and myocardial infarction (MI) size using late gadolinium enhancement (LGE), the efficacy of reperfusion therapy can be assessed as myocardial salvage in a single CMR session.

In the event of an acute coronary occlusion, a single artery is usually affected. As a consequence of the occlusion, transmural ischemia occurs within the affected

\* Correspondence: [einar.heiberg@med.lu.se](mailto:einar.heiberg@med.lu.se)

<sup>1</sup>Department of Clinical Physiology, Skåne University Hospital, Lund University, Lund, Sweden

Full list of author information is available at the end of the article

coronary artery's perfusion territory [5,6]. The myocardium subjected to ischemia becomes edematous and shows an increased signal intensity in T2-weighted CMR compared to non-ischemic myocardium [7]. Several techniques have been proposed for quantitative assessment of MaR in T2-weighted CMR, such as manual delineation [4], and threshold methods of two standard deviations (2SD) from remote [5,8], full width half maximum (FWHM) intensity [9] and Otsu [10]. Human observers delineating MaR take into account both regional intensity differences and *a priori* knowledge on perfusion territories and transmuralities, which may improve accuracy of MaR quantification. However, manual delineation is subjective and time consuming. Semi-automatic and automatic threshold methods such as 2SD, FWHM and Otsu have been proposed as more objective methods. A more advanced automatic algorithm for quantification of edema in T2-weighted CMR has recently been developed by Johnstone et al [11]. Their algorithm shows promising result for an automatic segmentation approach of edema and thereby MaR in T2-weighted CMR by incorporating regional analysis. However, neither the threshold methods, (2SD, FWHM and Otsu), nor the algorithm by Johnstone et al. uses *a priori* knowledge on the appearance of MaR and the cardiac anatomy, which is considered when performing manual delineations.

Therefore, the aim of this study was to develop an automatic segmentation algorithm for quantification of MaR in T2-weighted CMR images which uses *a priori* knowledge on the appearance of MaR and cardiac anatomy.

## Methods

### Study Population

The study was approved by the local ethics committee and all patients gave their written informed consent. Forty seven patients (age  $60.3 \pm 9.8$  years, range 39 - 83, 39 males) with first-time acute ST-elevation myocardial infarction (STEMI) due to a single occluded coronary artery confirmed by angiography were prospectively included in the study. All patients were treated with primary percutaneous coronary intervention (PCI) with coronary stenting, resulting in TIMI grade 3 flow in the culprit artery.

### CMR imaging

Within a week after admission patients were imaged in the supine position using either a 1.5 T system (Magnetom Vision, Siemens, Erlangen, Germany) with a CP body array coil or a 1.5 T system (Philips Intera CV or Achieva, Philips, Best, the Netherlands) with a cardiac synergy coil. Initial scout images were acquired to locate the heart, and a T2-weighted triple inversion turbo spin

echo sequence (STIR) was employed to depict the myocardium at risk. T2-weighted CMR images were acquired in the short-axis view, covering the left ventricle from the base to apex. Imaging parameters were: echo time 43 ms (Siemens) or 100 ms (Philips); repetition time 2 heart beats; number of averages 2; inversion time 180 ms; typical image resolution  $1.5 \times 1.5$  mm (Siemens) or  $1.4 \times 1.7$  mm reconstructed to  $0.7 \times 0.7$  mm (Philips); slice thickness 8 mm with a typical slice gap of 2 mm. When acquiring images with the cardiac synergy coil no parallel imaging was performed (SENSE = 1).

### Image analysis

The MaR was manually delineated according to the method previously described by Carlsson et al [4]. In short, endocardial and epicardial borders of the LV were traced in all short-axis slices by an experienced observer and the papillaries were excluded from the myocardium. Regions of hyper enhanced myocardium were manually delineated as myocardium at risk (MaR) by an experienced observer and expressed as percent of left ventricular mass (LVM). Hypo-intense myocardium within the area of increased signal intensity was regarded as microvascular obstruction [12] and was included in the MaR.

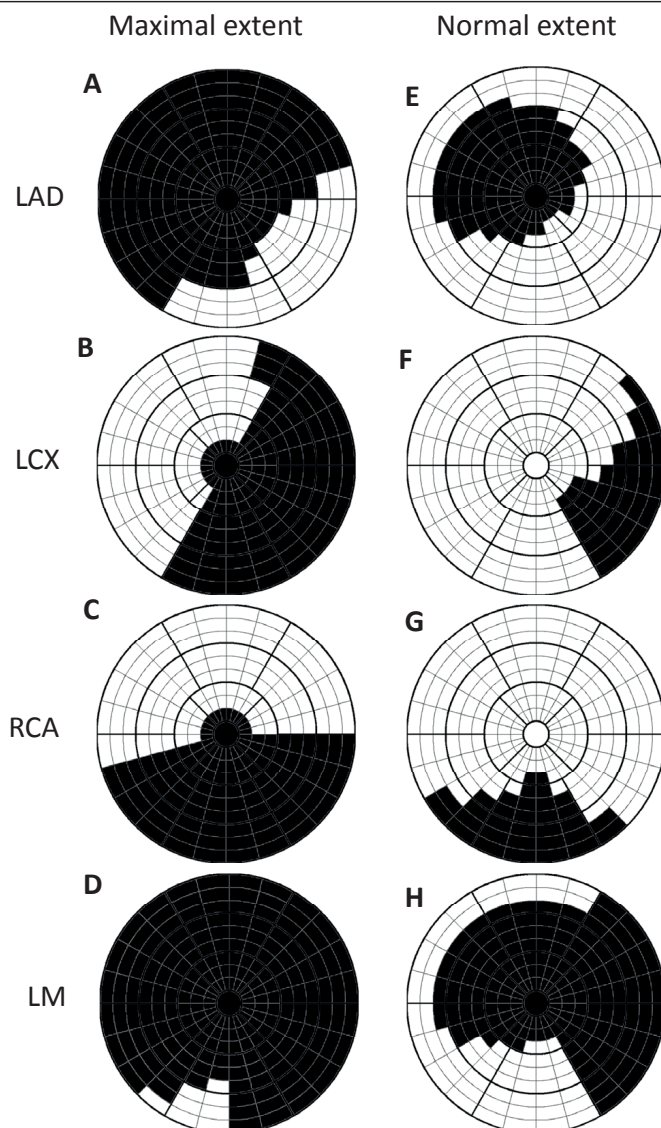
The new segmentation algorithm, called Segment MaR, was implemented in the freely available cardiac image analysis software Segment (<http://segment.hei-berg.se>) [13] and will be made available at time of publication. Segment was also used for manual delineation and implementation of the threshold methods (2SD, FWHM and Otsu).

### Automatic segmentation algorithm, Segment MaR

The automatic segmentation algorithm, Segment MaR, defines the MaR within the manually delineated left ventricular myocardium based on the culprit artery defined by the user. The MaR region is defined as a continuous region which has a higher probability of being MaR compared to normal myocardium, based on the signal intensity, and fulfills *a priori* criteria for MaR regarding transmuralities, shape, size and extent within the perfusion territory of the culprit artery.

Figure 1 shows a model of normal and maximal extent for the perfusion territories of each coronary artery as defined in consensus by three experienced observers from their combined experience of CMR and SPECT. Normal and maximal extent models were defined by each observer and discussed until consensus was reached for left anterior descending artery (LAD), left circumflex artery (LCx) and right coronary artery (RCA). The models for left main artery (LM) were defined from the models of LAD, LCx and RCA. The





**Figure 1 Maximal and normal extent model.** Bulls-eye representation of maximal extent model (left column) and normal extent model (right column) for the perfusion territories of left anterior descending artery (LAD), left circumflex artery (LCx), right coronary artery (RCA), and left main artery (LM). Models for LAD, LCx and RCA were defined in consensus by three experienced observers in an extended 17-segment AHA model and models for LM were defined from the models of LAD, LCx and RCA. The 17-segment model is extended to three slices in each of the basal, mid-ventricular and apical zones and 24 sectors in each slice. Black sectors are included in the maximal and normal extent model, respectively. The septal part of the left ventricle is represented in the left of the bulls-eye plot, the lateral part in the right, anterior part in the top, inferior part in the bottom, the apical slices in the center and the basal slices in the outer part of the bulls-eye plot.

extent was defined in the 17 segment AHA model [14] which was extended to 24 sectors circumferentially and 3 slices in each of the basal, mid-ventricular and apical parts of the left ventricle. The maximal and normal extent model of the user defined culprit artery is used as *a priori* information in the algorithm.

The Segment MaR algorithm can be divided into 7 steps (Figure 2)

- 1) User input as culprit artery and orientation
- 2) Estimate intensity distribution for normal myocardium and MaR
- 3) Define MaR probability from intensity distributions
- 4) Calculate MaR probability in 24 sectors for each short axis slice
- 5) Find region of connected sectors with high probability of MaR which fulfills regional criteria
- 6) Interpolate shape of MaR region over slices by normalized averaging
- 7) Define MaR region in the short-axis view from sector based segmentation

In step 1) the user defines the culprit artery as either, LAD, LCx, RCA or LM based on the overall appearance of the hyperenhanced region and indicates the orientation of the heart by indicating the inferior insertion point for the right ventricle (Figure 2:1).

In step 2) the intensity distribution for normal myocardium and MaR is estimated from the intensity histogram by an expectation maximization (EM) algorithm slice by slice [15]. The intensity distributions were analyzed slice by slice since the intensity can vary between slices in T2-weighted CMR images. The EM-algorithm estimates the mean and standard deviation of the intensity distributions for MaR and normal myocardium by refining an initial estimation. The initial estimation of mean and standard deviation for MaR is calculated from the intensities within the normal extent model for the culprit artery and for normal myocardium calculated from the intensities outside the maximal extent model for the culprit artery. The intensity distributions for MaR and normal myocardium are assumed to be Gaussian and are thereby defined by their mean and standard deviation as estimated by the EM-algorithm. Figure 2:2 shows the intensity histogram of the myocardium and the estimated Gaussian intensity distributions for normal myocardium and MaR.

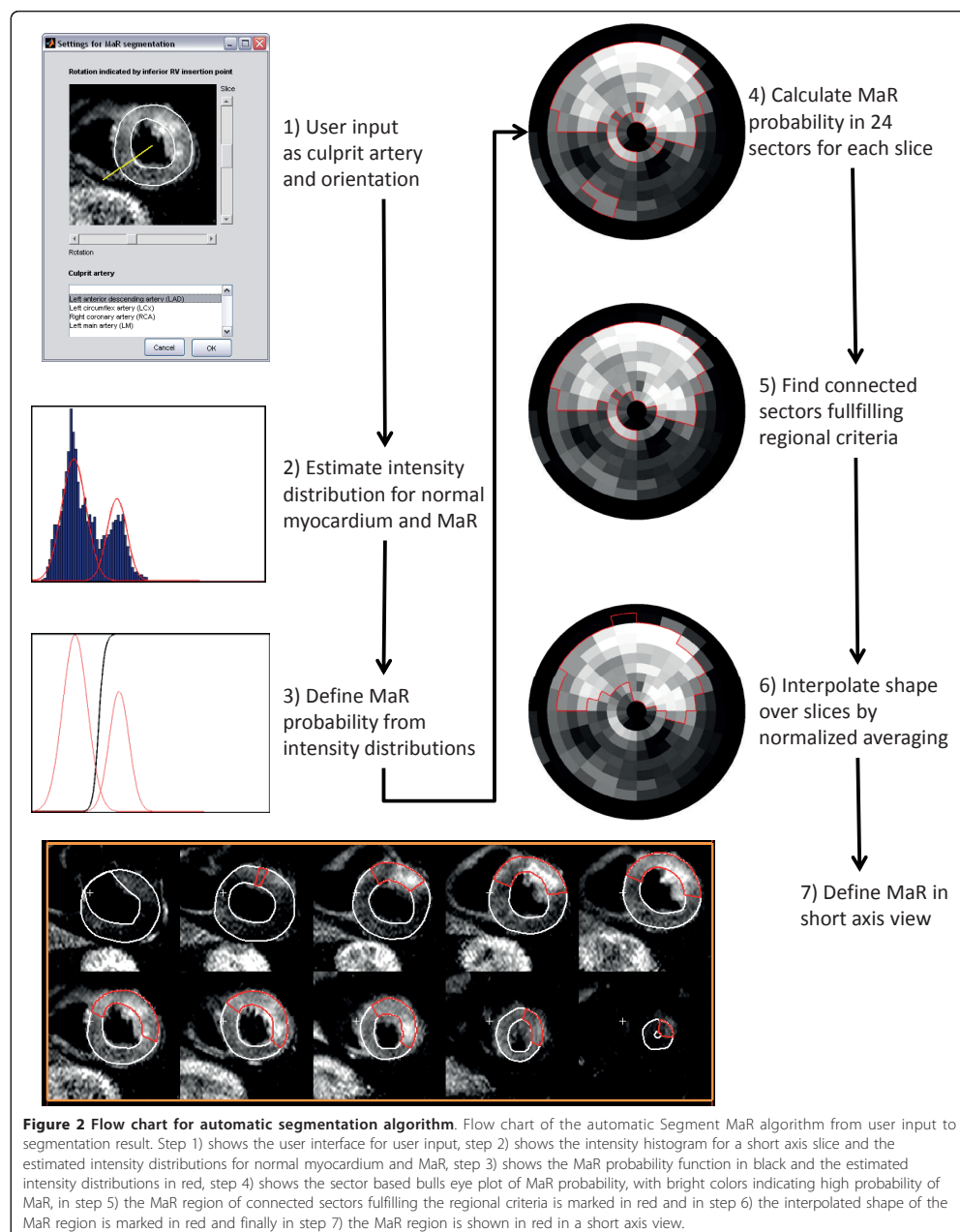
In step 3) a MaR probability is defined from the intensity distributions for normal myocardium and MaR. The MaR probability is defined from the Bayesian probability of MaR given that the intensity is either MaR or normal myocardium by dividing the Gaussian intensity distribution for MaR by the sum of the Gaussian distribution

for MaR and normal myocardium. Thus, the probability function is in the range from 0 to 1 and values above 0.5 indicate higher probability of MaR than normal myocardium. A probability function and its corresponding Gaussian distributions are shown in Figure 2:3.

In step 4) the MaR probability is calculated for 24 sectors in each short axis slice. The probability function in step 3) is defined for each short axis slice and mapped to each pixel in the slice. The probability value is then averaged for 24 sectors in each slice resulting in a sector-based MaR probability. The sector-based MaR probability is shown in Figure 2:4 where bright colors indicate high probability of MaR and the red border indicates sectors with a probability value above 0.5.

In step 5) a region of connected sectors with a high probability for MaR which fulfills regional criteria is identified. The criteria to be fulfilled is a) sectors with a probability value above 0.5 should be connected to its nearest neighboring sector within the slice or in an adjacent slice in a 4-neighbourhood to constitute a region, b) sectors should be localized within the maximal extent model for the culprit artery, c) in the slices with outflow tract only sectors on the anterior side of the outflow tract is considered MaR for LAD and LM and only sectors on the inferior side for LCx and RCA. Finally, the MaR probability of each region is calculated by summing the probability value of being MaR for each pixel within the region and summing the probability of being normal i.e. 1 minus the probability of being MaR for each pixel outside the region. If multiple connected regions are found, the region with the highest probability is chosen and the other regions are eliminated from the MaR region. The outer boundary of the new MaR region is indicated with a red border in Figure 2:5.

In step 6), the shape of the MaR region is refined by interpolating the outer boundary of the MaR region over slices by normalized averaging [16]. The normalized averaging interpolates the outer boundary by using certainty values for the outer boundary of each slice and a narrow kernel with width of 3 slices in both apical and basal direction. The certainty value is lowered if a) the region is close to the maximal extent model, b) if the difference in extent deviates from normal difference between slices and c) if the intensity appearance does not match the boundary of the MaR region. The certainty based on closeness to the maximal extent, a), is calculated by a linear function from one to zero between the normal and maximal extent model in each slice. The certainty based on difference in extent between slices, b), is calculated from a Gaussian function with standard deviation of two sectors and a mean of increasing two sectors from base to apex for LAD and LM respectively a mean of decreasing one sector for RCA and LCx. The certainty based on intensity appearance, c), was



**Figure 2 Flow chart for automatic segmentation algorithm.** Flow chart of the automatic Segment MaR algorithm from user input to segmentation result. Step 1) shows the user interface for user input, step 2) shows the intensity histogram for a short axis slice and the estimated intensity distributions for normal myocardium and MaR, step 3) shows the MaR probability function in black and the estimated intensity distributions in red, step 4) shows the sector based bulls eye plot of MaR probability, with bright colors indicating high probability of MaR, in step 5) the MaR region of connected sectors fulfilling the regional criteria is marked in red and in step 6) the interpolated shape of the MaR region is marked in red and finally in step 7) the MaR region is shown in red in a short axis view.

calculated as the MaR probability, as defined in step 5, for each slice. The new boundary defined by the normalized averaging is interpolated over slices to give a smooth appearance of the MaR region (Figure 2:6).

In step 7) the MaR region is defined in a short axis view from the sector based MaR region. This is done by defining all pixels within a sector as MaR if the sector is within the outer boundary of the sector based MaR region (Figure 2:7).

#### Comparison to other segmentation methods

The new automatic segmentation method, Segment MaR, was compared to inter observer variability of manual delineation and three threshold methods, 2 standard deviations from remote (2SD), full width half maximum intensity (FWHM) and Otsu. All methods used the same manual delineation of endocardium and epicardium and were applied slice by slice. A slice by slice approach was chosen since the intensity varies between slices in T2-weighted images.

A second experienced and independent observer manually delineated MaR in all subjects for inter observer analysis. The 2SD threshold method estimates an intensity threshold from a remote region as the mean plus two standard deviations of the intensity within the remote region. The remote region was defined as the region outside the maximal extent model of the culprit artery, indicated by the white sectors in the maximal extent model in Figure 1. The FWHM threshold method [17] estimates an intensity threshold from a remote region as midway between the mean intensity within the remote region and the maximal intensity within the myocardium. The remote region was defined in the same way as for 2SD. The threshold method of Otsu [18] estimates the intensity threshold from the histogram of all intensities to get minimal variance both above and below the threshold. The intensity threshold was calculated and applied slice by slice for all three threshold methods.

#### Statistical analysis

The quantification of MaR by the automatic Segment MaR algorithm, manual second observer delineation, the

threshold methods of 2SD, FWHM and Otsu were all compared against the reference observer using Bland-Altman bias (mean  $\pm$  standard deviation), paired t-test and linear regression analysis (correlation coefficient and p-value). Regional agreement to manual delineation by the reference observer was evaluated by calculating the Dice similarity coefficient (DSC) [19], which can be derived from the kappa statistics for the classification of pixels [20]. The DSC is calculated as two times the volume of the intersection of the MaR regions divided by the sum of the volumes of the two MaR regions. The DSC is therefore 0 if the regions do not overlap and 1 if the regions overlap perfectly. The DSC was calculated against the reference observer, for Segment MaR, second observer delineation, 2SD, FWHM and Otsu, for each patient and expressed as mean  $\pm$  standard deviation.

#### Results

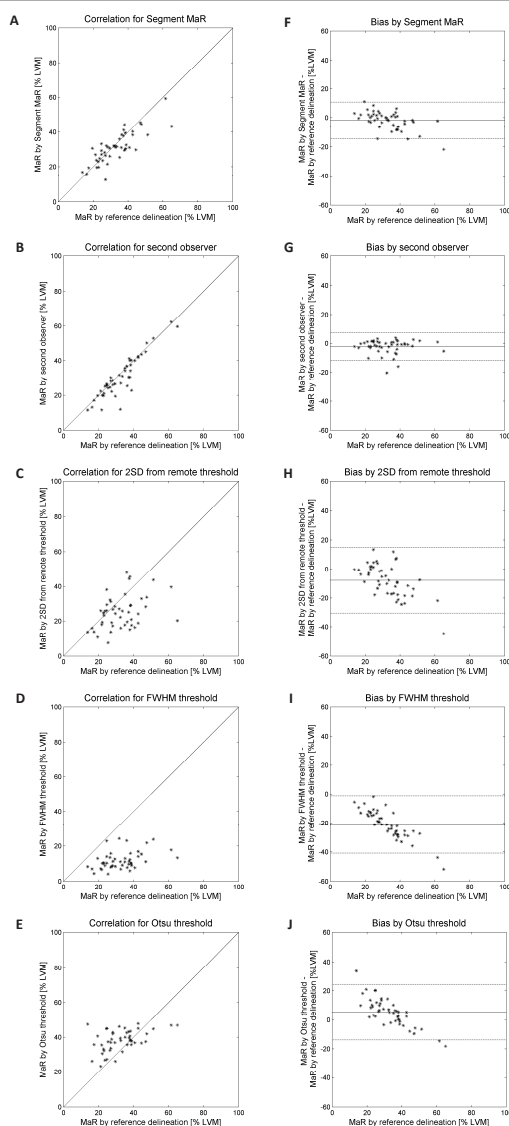
MaR assessed by the reference observer was  $32.9 \pm 10.9\%$  of LVM and MaR assessed by Segment MaR was  $31.0 \pm 8.8\%$ . There was a strong correlation,  $R = 0.81$ ,  $p < 0.001$ , and low bias,  $-1.9 \pm 6.4\%$  of LVM,  $p = 0.047$ , when Segment MaR was compared to the reference delineation of MaR (Table 1, Figure 3). The inter observer variability of manual delineation as the bias between reference and second observer was  $-2.3 \pm 4.9\%$  of LVM. The bias for Segment MaR was lower than for the threshold methods of 2SD, FWHM and Otsu,  $-7.7 \pm 11.4\%$  of LVM,  $-21.0 \pm 9.9\%$  of LVM and  $5.3 \pm 9.6\%$  of LVM, respectively (Table 1). Furthermore there was a good regional agreement between Segment MaR and the manual reference delineation,  $DSC = 0.85 \pm 0.07$  (Table 1). In Figure 4 typical segmentations for all five methods are shown in the same patient and compared to manual delineation by the reference observer. For Segment MaR and manual delineation by the reference and second observer, the MaR region is continuous whereas the segmentation by the threshold methods of 2SD, FWHM and Otsu consist of multiple regions of hyperenhanced myocardium.

In order to analyze the added value of each step in the Segment MaR algorithm (Figure 2) the bias to manual delineation by the reference observer was calculated

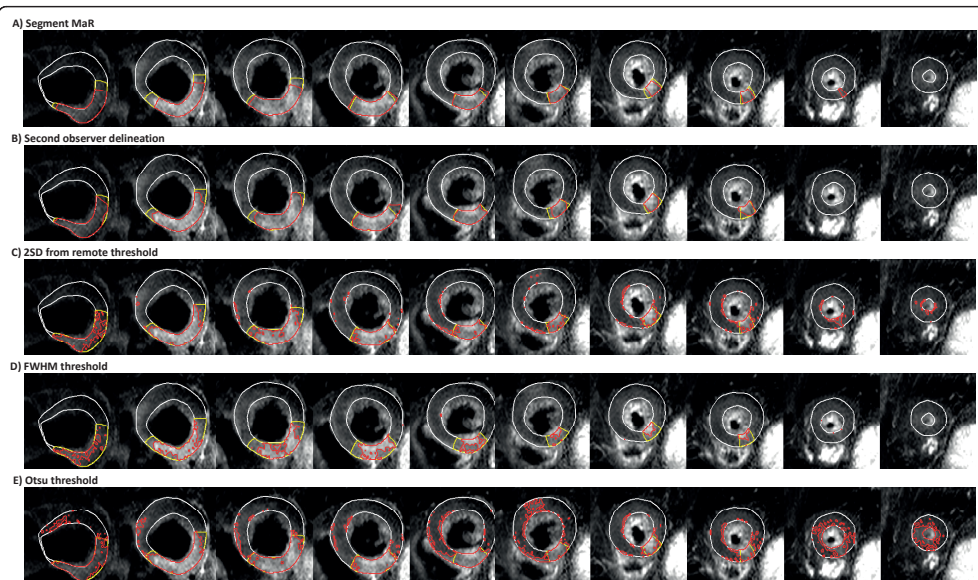
**Table 1 Results for all five segmentation methods compared to reference delineation**

	MaR bias [% of LVM]	p-value	Regression R-value	p-value	DSC
Segment MaR	$-1.9 \pm 6.4$	0.047	0.81	<0.001	$0.85 \pm 0.07$
Second observer delineation	$-2.3 \pm 4.9$	0.003	0.91	<0.001	$0.90 \pm 0.08$
2SD threshold	$-7.7 \pm 11.4$	<0.001	0.38	0.008	$0.69 \pm 0.14$
FWHM threshold	$-21.0 \pm 9.9$	<0.001	0.41	0.004	$0.46 \pm 0.14$
Otsu threshold	$5.3 \pm 9.6$	<0.001	0.47	<0.001	$0.68 \pm 0.10$

MaR-Myocardium at risk, DSC-Dice similarity coefficient, 2SD-two standard deviations from remote, FWHM-full width half maximum intensity



**Figure 3 Correlation and Bland-Altman plot for all five segmentation methods.** Correlation of MaR as % of LVM (panel A-E) and Bland-Altman plot of MaR bias as % of LVM (panel F-J) against the reference delineation for the automatic Segment MaR algorithm (panel A, F), second observer delineation (panel B, G), the threshold methods of two standard deviations from remote threshold (2SD)(panel C, H), full width half maximum (FWHM)(panel D, I) and Otsu (panel E, J).



**Figure 4 Typical segmentation result for all five segmentation methods.** Typical MaR segmentation shown in red for the automatic segmentation Segment MaR (panel A), second observer delineation (panel B), the threshold methods of two standard deviations from remote (2SD) (panel C), full width half maximum (FWHM) (panel D) and Otsu (panel E), compared to manual delineation by the reference observer, shown in yellow, all within the manual delineation of myocardial borders (shown in white). The same patient, short-axis slices, manual delineation of myocardial borders and manual reference delineation of MaR is used for all methods and shown from most basal slice in the left of the panel to the most apical slice in the right of the panel. Note the continuous appearance of the segmentation for Segment MaR and manual delineation by the reference and second observer compared to the threshold methods of 2SD, FWHM and Otsu.

after steps 3, 4, 5 and 6. As a base-line, bias was calculated for segmentation by the EM-algorithm without *a priori* information. The results are shown as mean  $\pm$  standard deviation in Figure 5.

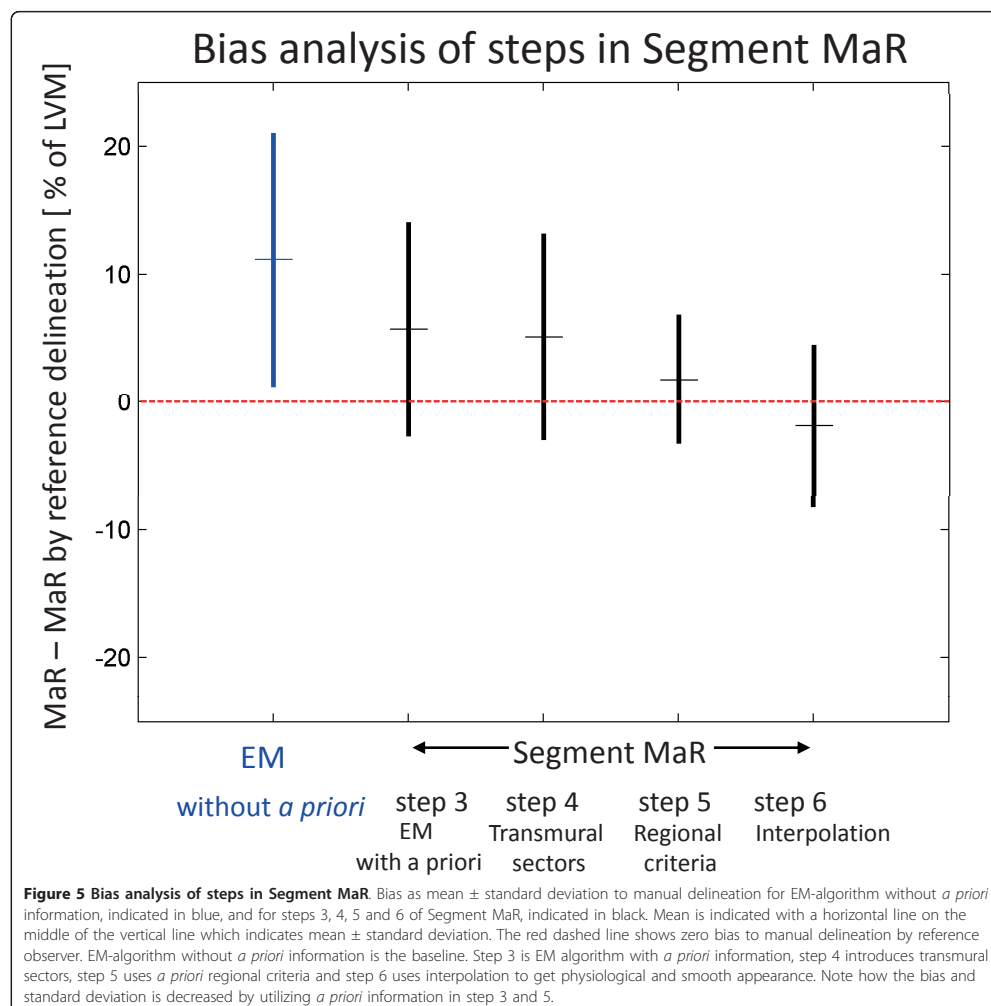
In eight of the forty-seven patients multiple hyperenhanced regions were detected in step 5 of the algorithm. In seven out of those the same region as by manual delineation was identified. In the one case, a dark artifact divided the MaR region into two disconnected regions, resulting in the Segment MaR algorithm only identifying one of these and subsequently underestimating the MaR region.

## Discussion

This study has presented an automatic segmentation algorithm for quantification of MaR from T2-weighted CMR, based on the EM-algorithm and *a priori* information on normal and maximal perfusion territories for the culprit artery. Compared to manual delineation by a reference observer, the new algorithm, Segment MaR, performed better than previously suggested threshold methods (2SD, FWHM and Otsu). Quantitative bias and

regional agreement for Segment MaR were similar to inter observer variability of manual delineation.

The new automatic segmentation algorithm, Segment MaR, estimates an intensity based probability of MaR from all intensity information in each short axis slice by an EM-algorithm. The use of all intensity information may make the estimate more robust to noise, artifacts and variation in signal homogeneity. The use of *a priori* information in the EM-algorithm showed an added value compared to using the EM-algorithm without *a priori* information in the bias analysis (Figure 5). The constraint of the extent by a maximal extent model eliminates artifacts located outside the perfusion territory and together with the use of only one region of connected sectors step 5 showed an added value in the bias analysis. Small non-transmural artifacts within the perfusion territory of the culprit artery may also be less likely to be considered as MaR by requiring transmural regions from the connected sectors. The bias analysis showed no added value by use of the interpolation although it is needed for a physiological appearance of the MaR region.



In the edema algorithm by Johnstone et al. [11] an EM-algorithm was used to estimate the intensity distributions and a threshold was defined as 2 standard deviations above the mean intensity of normal myocardium. Thereby the threshold is similar to that of 2SD threshold with the exception that the intensity of normal myocardium is defined from the EM-algorithm instead of a remote region. Using a threshold of 2SD does not utilize the intensity information on MaR which has been estimated by the EM-algorithm and thereby the variation in signal intensity may not be taken into consideration.

Another difference between Segment MaR and the edema algorithm by Johnstone is that the Segment MaR algorithm is based on 24 sectors in each slice instead of pixel wise segmentation. The pixel-wise segmentation does not consider transmuralities and may give better precision on the boundary of MaR. Pixel-wise segmentation may, however, be more sensitive to artifacts. Quantitative results for the edema algorithm, reported by Johnstone, were a bias of  $1.1 \pm 10.1\%$  of LVM and DSC of  $0.50 \pm 0.27$  to their reference of manual delineation and an aim for future work was set to reach  $DSC > 0.7$ .



which has been stated to indicate excellent regional agreement by Zijdenbos et al. [20]. In the present study, Segment MaR showed a DSC value of  $0.85 \pm 0.07$  and bias of  $-1.9 \pm 6.4\%$  of LVM compared to the manual delineation of the reference observer. Note-able is also the difference in DSC for inter observer variability of manual delineation which was  $0.72 \pm 0.14$  in the study by Johnstone compared to  $0.90 \pm 0.08$  in the present study. This may indicate that the regional agreement may be lower and the quantitative variability of manual delineation may be larger between observers than reported in the present study and thereby the use of an automatic algorithm such as Segment MaR, which has a low bias and high regional agreement to manual delineation, may decrease the inter observer variability.

The weak correlation to manual delineation by the reference observer for the threshold methods of 2SD, FWHM and Otsu may be explained by the fact that the methods are solely based on a threshold and by weaknesses in defining the threshold for the different methods. Using a fixed number of standard deviations as in 2SD does not account for the variability of intensity for MaR and more importantly the contrast in T2-weighted CMR is lower than for other CMR [21]. The threshold by FWHM may be sensitive to artifacts since the brightest pixel intensity is used to find the threshold. Both 2SD and FWHM are sensitive to the definition of the remote region which is currently not standardized. The remote region in this study was automatically defined as the myocardium outside the maximal extent model of the culprit artery. This definition of the remote myocardium was chosen to obtain an objective and standardized representation of normal myocardium. This strategy may, however, result in artifacts being included in the remote myocardium. An overestimation of the threshold may for 2SD be caused by the large standard deviation of the intensity within the remote region. The border zone between normal myocardium and remote myocardium influences the remote myocardium proportionally more for larger MaR regions which may result in an overestimation of the threshold and subsequent underestimation of the MaR region and thus explain the trend seen in bias for 2SD (Figure 3). For FWHM an overestimation of the threshold may be caused by artifacts within the myocardium and the remote region. Bright artifacts may result in a threshold which only identifies the artifact as the MaR region and this can explain the trend in bias seen for FWHM since it will result in a larger underestimation for larger MaR regions. The threshold defined by Otsu does not depend on any remote region but may instead be unstable in the definition of the threshold since it assumes that an optimal threshold should be found in each short axis slice. This implies that both MaR and normal

myocardium should be present in each slice, which is not the case in most patients as for example in the basal slices of an LAD occlusion or apical slices of an RCA occlusion. The Otsu threshold may thereby overestimate the MaR region in slices lacking MaR and underestimate MaR in slices lacking normal myocardium. This may explain the large overestimation for small MaR regions and large underestimation for large MaR regions. The EM-algorithm used in Segment MaR also assumes two intensity distributions as in the threshold method of Otsu but is accompanied by *a priori* information both as initialization to the EM-algorithm and in the post processing of finding connected sectors.

Manual monitoring and possibly manual corrections are as for all automatic segmentation algorithms needed for research and clinical use. The use of Segment MaR may, however, decrease the degree of variability introduced by the subjectivity of manual delineation since the Segment MaR algorithm showed a low bias and high correlation to manual delineation regarding quantitative assessment and an excellent regional agreement according to DSC.

One limitation in this study is the lack of ground truth for in vivo quantification of MaR. Manual delineation according to the methodology used by Carlsson et al. [4] when validating T2-weighted CMR for MaR to SPECT was chosen as reference method and in this study there was a good inter observer agreement. Due to the limited number of patients in the study it was not possible to use a separate training and test set and the parameters in the automatic Segment MaR algorithm could not be optimized. The Segment MaR algorithm has not been specifically designed for the imaging systems of Philips and Siemens. Further research is suggested to investigate the performance of the algorithm in a larger cohort of patients and possibly optimize and improve the algorithm for specific imaging systems.

## Conclusions

In this study, an automatic segmentation algorithm, called Segment MaR, for quantification of myocardium at risk (MaR) in T2-weighted CMR has been presented and showed to have a good agreement to manual delineation. Both the quantitative and regional agreement to manual delineation was better for Segment MaR than for the threshold methods of two standard deviations from remote, full width half maximum intensity and Otsu. The Segment MaR algorithm seems to be a promising, objective method for standardized measurement of MaR in T2-weighted CMR.

## Acknowledgements

The authors would like to acknowledge Ann-Helen Arvidsson and Christel Carlander, both with the Lund Cardiac MR Group, for their skillful assistance



with image acquisition. This study has been funded by the Swedish Research Council (2008-2461, 2008-2949), The Swedish Heart and Lung Foundation, The Medical Faculty of Lund University, Sweden, and Region of Scania, Sweden.

# Author details

<sup>1</sup>Department of Clinical Physiology, Skåne University Hospital, Lund University, Lund, Sweden. <sup>2</sup>Department of Numerical Analysis, Centre for Mathematical Sciences, Lund University, Lund, Sweden.

# Authors' contributions

All authors contributed to concept and design of the study, input to the development of an automatic segmentation algorithm and revised the manuscript. All authors read and approved the final manuscript. JS developed and implemented Segment MaR, implemented the threshold methods, analyzed and interpreted segmentation results, and drafted the manuscript. JU included patients to the study and performed manual delineation of the left ventricle. HE performed manual delineation of MaR as second observer, defined maximal and normal extent models, and analyzed and interpreted data. MC defined maximal and normal extent models, and analyzed and interpreted data. HA performed manual delineation of MaR as reference observer, defined maximal and normal extent models and conceived the study. EH analyzed and interpreted data and conceived the study.

# Competing interests

EH is the founder of Medviso AB, Lund, Sweden, which sells a commercial version of Segment. JS is employed by Medviso AB on a part-time basis. The other authors declare that they have no competing interests.

Received: 23 August 2011 Accepted: 31 January 2012

Published: 31 January 2012

# References

1. Aletras AH, Tilak GS, Natanzon A, Hsu LY, Gonzalez FM, Hoyt RF Jr, Arai AE. Retrospective determination of the area at risk for reperfused acute myocardial infarction with T2-weighted cardiac magnetic resonance imaging: histopathological and displacement encoding with stimulated echoes (DENSE) functional validations. *Circulation*. 2006;**113**:1865–1870.
2. Fishbein MC, Maclean D, Maroko PR. The histopathologic evolution of myocardial infarction. *Chest*. 1978;**73**:843–849.
3. Jennings RB, Schaper J, Hill ML, Steenberg C Jr, Reimer KA. Effect of reperfusion late in the phase of reversible ischemic injury. Changes in cell volume, electrolytes, metabolites, and ultrastructure. *Circ Res*. 1985;**56**:262–278.
4. Carlsson M, Ubachs JF, Hedstrom E, Heiberg E, Jovinge S, Arheden H. Myocardium at risk after acute infarction in humans on cardiac magnetic resonance: quantitative assessment during follow-up and validation with single-photon emission computed tomography. *JACC Cardiovasc Imaging*. 2009;**2**:569–576.
5. Friedrich MG, Abdel-Aty H, Taylor A, Schulz-Menger J, Messroghli D, Dietz R. The salvaged area at risk in reperfused acute myocardial infarction as visualized by cardiovascular magnetic resonance. *J Am Coll Cardiol*. 2008;**51**:1581–1587.
6. Stork A, Muellerleile K, Bansmann PM, Graessner J, Kaul M, Kemper J, Adam G, Lund GK. Value of T2-weighted, first-pass and delayed enhancement, and cine CMR to differentiate between acute and chronic myocardial infarction. *Eur Radiol*. 2007;**17**:610–617.
7. Abdel-Aty H, Zagrosek A, Schulz-Menger J, Taylor AJ, Messroghli D, Kumar A, Gross M, Dietz R, Friedrich MG. Delayed enhancement and T2-weighted cardiovascular magnetic resonance imaging differentiate acute from chronic myocardial infarction. *Circulation*. 2004;**109**:2411–2416.
8. Wright J, Adriaenssens T, Dymarkowski S, Desmet W, Bogaert J. Quantification of myocardial area at risk with T2-weighted CMR: comparison with contrast-enhanced CMR and coronary angiography. *JACC Cardiovasc Imaging*. 2009;**2**:825–831.
9. Tilak GS, Hsu LY, Hoyt RF Jr, Arai AE, Aletras AH. In vivo T2-weighted magnetic resonance imaging can accurately determine the ischemic area at risk for 2-day-old nonreperfused myocardial infarction. *Invest Radiol*. 2008;**43**:7–15.
10. Burchell T, Flett A, Petersen S, Davies L, Mohiddin S, Mathur A, Westwood M. Comparing analysis methods for quantification of myocardial oedema in patients following reperfused ST-elevation MI. *Journal of Cardiovascular Magnetic Resonance*. 2011;**13**:M11.
11. Johnstone RI, Greenwood JP, Biglands JD, Plein S, Ridgway JP, Radjenovic A. Assessment of tissue edema in patients with acute myocardial infarction by computer-assisted quantification of triple inversion recovery prepared MRI of the myocardium. *Magn Reson Med*. 2011.
12. Beek AM, Nijveldt R, van Rossum AC. Intramyocardial hemorrhage and microvascular obstruction after primary percutaneous coronary intervention. *Int J Cardiovasc Imaging*. 2010;**26**:49–55.
13. Heiberg E, Sjogren J, Ugander M, Carlsson M, Engblom H, Arheden H. Design and validation of Segment—freely available software for cardiovascular image analysis. *BMC Med Imaging*. 2010;**10**:1.
14. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, Pennell DJ, Rumberger JA, Ryan T, Verani MS. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Circulation*. 2002;**105**:539–542.
15. Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data Via Em Algorithm. *J Roy Stat Soc B Met*. 1977;**39**:1–38.
16. Granlund G, Knutsson H. Signal processing for computer vision. Linköping: Kluwer Academic Publishers; 1995.
17. Hsu LY, Natanzon A, Kellman P, Hirsch GA, Aletras AH, Arai AE. Quantitative myocardial infarction on delayed enhancement MRI. Part I: Animal validation of an automated feature analysis and combined thresholding infarct sizing algorithm. *J Magn Reson Imaging*. 2006;**23**:298–308.
18. Otsu N. Threshold Selection Method from Gray-Level Histograms. *IEEE T Syst Man Cyb*. 1979;**9**:62–66.
19. Dice LR. Measures of the Amount of Ecologic Association between Species. *Ecology*. 1945;**26**:297–302.
20. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric Analysis of White-Matter Lesions in Mr-Images-Method and Validation. *IEEE T Med Imaging*. 1994;**13**:716–724.
21. Abdel-Aty H, Simonetti O, Friedrich MG. T2-weighted cardiovascular magnetic resonance imaging. *J Magn Reson Imaging*. 2007;**26**:452–459.

doi:10.1186/1532-429X-14-10

Cite this article as: Sjögren et al.: Semi-automatic segmentation of myocardium at risk in T2-weighted cardiovascular magnetic resonance. *Journal of Cardiovascular Magnetic Resonance* 2012 **14**:10.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit





## Paper IV



# Automatic segmentation of myocardium at risk from contrast enhanced SSFP CMR: validation against expert readers and SPECT

Jane Tufvesson<sup>1,2</sup>, Marcus Carlsson<sup>1</sup>, Anthony H. Aletras<sup>1,3</sup>, Henrik Engblom<sup>1</sup>, Jean-François Deux<sup>4</sup>, Sasha Kouf<sup>5</sup>, Peder Sörensson<sup>6</sup>, John Pernow<sup>6</sup>, Dan Atar<sup>7</sup>, David Erlinge<sup>5</sup>, Håkan Arheden<sup>1</sup>, Einar Heiberg<sup>1,2</sup>

## Background

Efficacy of reperfusion therapy can be assessed as myocardial salvage index (MSI) by determining the size of myocardium at risk (MaR) and myocardial infarction (MI), ( $MSI = 1 - MI/MaR$ ). Cardiovascular magnetic resonance (CMR) can be used to assess MI by late gadolinium enhancement (LGE) and MaR by either T2-weighted imaging or contrast enhanced SSFP (CE-SSFP). Automatic segmentation algorithms have been developed and validated for MI by LGE as well as for MaR by T2-weighted imaging. There are, however, no algorithms available for CE-SSFP. Therefore, the aim of this study was to develop and validate automatic segmentation of MaR in CE-SSFP.

## Methods

The automatic algorithm applies surface coil intensity correction and classifies myocardial intensities by Expectation Maximization to define a MaR region based on *a priori* regional criteria, and infarct region from LGE. Automatic segmentation was validated against manual delineation by expert readers in 183 patients with reperfused acute MI from two multi-center randomized clinical trials (RCT) (CHILL-MI and MITOCARE) and against myocardial perfusion SPECT in an additional set ( $n=16$ ). Endocardial and epicardial borders were manually delineated at end-diastole and end-systole. Manual delineation of MaR was used as reference and inter-observer variability was assessed for both manual delineation and automatic segmentation of MaR in a subset of patients ( $n=15$ ). MaR was expressed as percent of left ventricular mass (%LVM) and analyzed by bias ( $\text{mean} \pm \text{standard deviation}$ ). Regional agreement was analyzed by Dice Similarity Coefficient (DSC) ( $\text{mean} \pm \text{standard deviation}$ ).

## Results

MaR assessed by manual and automatic segmentation were  $36 \pm 10\%$  and  $37 \pm 11\%$  LVM respectively with bias  $1 \pm 6\%$  LVM and regional agreement DSC  $0.85 \pm 0.08$  ( $n=183$ ). MaR assessed by SPECT and CE-SSFP automatic segmentation were  $27 \pm 10\%$  LVM and  $29 \pm 7\%$  LVM respectively with bias  $2 \pm 7\%$  LVM. Inter-observer variability was  $0 \pm 3\%$  LVM for manual delineation and  $-1 \pm 2\%$  LVM for automatic segmentation.

## Conclusions

Automatic segmentation of MaR in CE-SSFP was validated against manual delineation in multi-center, multi-vendor studies with low bias and high regional agreement. Bias and variability was similar to inter-observer variability of manual delineation and inter-observer variability was decreased by automatic segmentation. Thus, the proposed automatic segmentation can be used to reduce subjectivity in quantification of MaR in RCT.

## Background

Myocardium at risk (MaR) is defined as the ischemic myocardium during coronary artery occlusion, at risk of infarction if the blood flow in the occluded artery is not restored in time. The myocardial infarction evolves during time to treatment and if blood flow is not restored in time the whole region of MaR becomes myocardial infarction (MI). If both the size of MaR and final MI size is determined, the efficacy of reperfusion therapy can be assessed as myocardial salvage index ( $MSI = 1 - MI/MaR$ ). By using MSI instead of MI size alone the number of patients needed in clinical trials can be reduced [1] since MI size is related to MaR which is specific for each patient and coronary occlusion.

Cardiovascular magnetic resonance (CMR) is considered gold standard for assessment of infarct size by late gadolinium enhancement (LGE) [2]. Myocardial perfusion SPECT is considered gold standard for assessment of MaR but requires an radioactive isotope to be injected before the blood flow is restored in occluded artery and imaging is performed only hours after the treatment. By CMR MaR can be assessed by either T2-weighted imaging [3] or contrast enhanced steady state free precession (CE-SSFP) [4] and both have been validated against SPECT for assessment of MaR up to one week after MI [5] [4]. Recently, both T2-weighted imaging and CE-SSFP have been used to determine myocardial salvage in two multi-center cardioprotective studies, CHILL-MI [6] and MITOCARE [7]. In these multi-center trials CE-

SSFP was shown to provide significantly better diagnostic image quality than T2-weighted images and to be more robust across vendors [8]. CE-SSFP may therefore be more suitable than T2-weighted imaging for quantification of MaR in multi-center settings.

An automatic segmentation algorithm is preferable for objective quantification in order to reduce subjectivity as well as time required for image analysis. Several algorithms have been developed and validated for automatic segmentation of MI size in LGE images [9-11]. Two automatic algorithms have been developed and validated in T2-weighted images, one specifically for MaR [12] and one for edema [13]. However, no algorithm has been developed yet for quantification of MaR in CE-SSFP images. Automatic quantification of MaR in T2-weighted images has been shown to yield more accurate results when utilizing Expectation Maximization (EM) to classify myocardial intensities and adding an *a priori* model of the perfusion territories compared to thresholding methods such as two standard deviations (2SD) from remote, full width half maximum (FWHM) and Otsu's method for quantification of MaR in T2-weighted images [12]. Therefore, the aim of this study was to develop and validate this automatic segmentation algorithm for MaR in CE-SSFP.

## Methods

### Study population and design

For validation of the automatic algorithm, patients with first time ST-elevation myocardial infarction (STEMI) treated with percutaneous coronary intervention (PCI) who had undergone CMR examination with CE-SSFP and LGE images of diagnostic quality as a part of the recently published clinical cardioprotection trials CHILL-MI [6] (n=92) and MITOCARE [7] (n=91) were included (n=183). Patients underwent CMR imaging within 2-6 days following acute MI treated with PCI. Inclusion and exclusion criteria for each of the clinical trials have been previously published [6, 14]. In short, all patients had clinical signs of acute myocardial infarction defined as clinical symptoms and ECG signs consistent with ST-elevation infarction or new onset of left bundle branch block (LBBB), were  $\geq 18$  years old and had symptom duration of less than 6 hours. Patients with a history of previous myocardial infarction or history of coronary revascularization were excluded.

For validation against an independent reference method of imaging MaR, an additional set of patients who had undergone both CE-SSFP CMR and single photon emission computed tomography (SPECT) (n=16) [4] were included in this study.

Inclusion and exclusion criteria for this cohort have also been previously published [4]. In short, all patients had clinical signs of acute myocardial infarction defined as clinical symptoms and ECG signs consistent with ST-elevation infarction and chest pain  $\geq 30$  minutes and  $\leq 9$  hours. Patients with a history of previous myocardial infarction or history of coronary revascularization were excluded.

All three studies [4, 6-7] from which patients were included were approved by the institutional review boards/ethics committees, and all patients provided written informed consent. No specific ethics approval or informed consent was needed for the development of the new automatic algorithm in the current study.

### Imaging

All CMR examinations were performed on 1.5 T scanners from Philips (Philips Healthcare, Best, The Netherlands), Siemens (Siemens AG, Erlangen, Germany) or GE (GE Healthcare, Waukesha, WI, USA). For visualization of MaR and evaluation of left ventricular function, CE-SSFP cine images were obtained approximately 5 minutes after intravenous injection of 0.2 mmol per kilogram of body weight of an extracellular gadolinium-based contrast agent [4, 6, 14]. The slice thickness was 8 mm with no slice gap. In-plane resolution was typically 1.5 x 1.5 mm. Typically, 20-30 CE-SSFP images were acquired per cardiac cycle. For infarct visualization LGE images covering the entire left ventricle were acquired approximately 15 minutes after injection of the gadolinium-based contrast agent. The LGE-images were acquired using an inversion-recovery gradient-recalled echo sequence with a slice thickness of 8 mm with no slice gap [15]. In-plane resolution was typically 1.5 x 1.5 mm. Inversion time was manually adjusted to null the signal of viable myocardium. Surface coil intensity correction was not generally applied across vendors and sites.

SPECT was performed in the additional set of 16 patients. Prior to opening the occluded vessel an intravenous injection of  $^{99m}\text{Tc}$ -tetrofosmin body weight adjusted (350-700 MBq) was administered to the patient. Myocardial perfusion SPECT imaging was performed within four hours to visualize and quantify MaR using either of two dual head cameras: GE (Venti, GE Healthcare, Waukesha, WI, USA) or Sopha (DST-XL, Sopha Medical Vision, Bue, Cedex, France). Typical pixel size was 6.4 x 6.4 x 6.4 mm (GE) and 3 x 3 x 3 mm (Sopha). Short axis images were reconstructed semi-automatically on the workstation for each camera.

## Image analysis

Both CMR and SPECT images were analyzed using the software Segment (<http://segment.heiberg.se>) [16].

In CE-SSFP images, MaR was manually assessed from short-axis images according to previously described methods [4, 6-7]. In short, the left ventricular myocardium was defined by manually delineating the epicardial and endocardial borders both at end-diastole and at end-systole as previously described. Hyper-intense regions within the myocardium in CE-SSFP images were manually delineated for assessment of MaR. Hypo-intense myocardium within the area of increased signal intensity was regarded as microvascular obstruction [17] and was included in the MaR. The delineation of each data set was performed by one of three primary observers with a quality control of the delineations by a second opinion for each case. Different opinions for the delineation were resolved in consensus between all three observers when necessary. All three observers had long experience in the field of CMR (HE, MC and HA with 14, 15 and 20 years of experience, respectively). MaR was expressed as percent of left ventricular mass (%LVM) [18]. In a subset of 15 patients from the multi-center studies, second observer analysis was performed to evaluate inter-observer variability (MC vs. HE).

In LGE images, infarct was delineated from the short-axis images according to a previously validated method [9]. In short, the endocardial and epicardial borders were traced manually with exclusion of the papillary muscles. The LGE myocardium was defined using a previously validated automatic segmentation algorithm [9] which is based on a 1.8SD from remote threshold, region analysis and a weighted summation according to pixel intensities to take partial volume effects into consideration. Manual adjustments were made when obvious image artefacts caused misinterpretation by the automatic algorithm and to include micro vascular obstruction when not detected by the algorithm. Hypointense regions within the region of LGE as a sign of microvascular obstruction [17], were included in the analysis as 100 % infarction.

In SPECT images, MaR was delineated by use of an 55 % threshold [19] and manual corrections after automatic delineation of epicardial and endocardial borders [20]. MaR was expressed as percent of left ventricular mass (%LVM).

Image quality was manually assessed as (1) non-diagnostic, (2) acceptable or (3) good. Acceptable and good images were considered to be of diagnostic quality and only CE-SSFP images with

diagnostic quality and full coverage of the left ventricle were included in this study as test set (n=183, Supplemental figure 1) and additional set (n=16). Patient characteristics of the test set are reported in **Table 1**.

**Table 1 Patient characteristics from test set n=183**

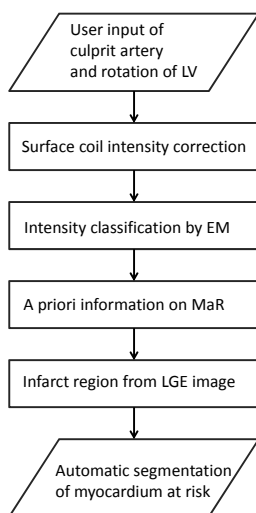
	medel $\pm$ SD	(min,max)
HR[beats/min]	68 $\pm$ 12	(31, 111)
EDV [ml]	178 $\pm$ 43	(32, 336)
ESV [ml]	94 $\pm$ 32	(20, 240)
EF [%]	48 $\pm$ 9	(19, 70)
LVM [g]	124 $\pm$ 28	(25, 252)
IS [%LVM]	17 $\pm$ 10	(2, 47)
MVO [%LVM]	3 $\pm$ 5	(0, 27)

*HR: heart rate, EDV: end-diastolic volume, ESV: end-systolic volume, EF: ejection fraction, LVM: left ventricular mass, IS: infarct size, MVO: micorvascular obstruction*

## Automatic segmentation algorithm

The automatic segmentation algorithm was originally developed for segmentation of MaR in T2-weighted images [12] and has in this study been developed for CE-SSFP images. Maximal extent models of perfusion territories for each coronary artery [12] were defined by expert observers and used to define remote and culprit region. The maximal extent models correspond to the MaR region of proximal occlusions and takes anatomy variations between patients into consideration. As input to the automatic algorithm, the manual delineation of endocardial and epicardial borders is used and the user defines the culprit artery as either left anterior descending artery (LAD), left circumflex artery (LCx), right coronary artery (RCA), or left main artery (LM) based on the overall appearance of the hyper enhanced region and defines right ventricular insertion points in CE-SSFP and LGE images, to find how to rotate the maximal extent model.

The automatic algorithm consist of four processing blocks after user input as shown in **Figure 1**, 1) surface coil intensity correction, 2) classification of myocardial intensities by Expectation Maximization (EM) [21], 3) definition of MaR region based on *a priori* regional criteria, and 4) incorporation of infarct region from LGE images.



**Figure 1 Automatic segmentation algorithm**

*The new automatic algorithm for segmentation of myocardium at risk (MaR) in CE-SSFP lets the user define the culprit artery and the rotation of the left ventricle as input. The algorithm consists of four processing blocks, surface coil intensity correction, intensity classification by Expectation Maximization (EM), segmentation based on a priori information on MaR and incorporation of infarct region from LGE images.*

Surface coil intensity correction is applied as a second order linear correction based on the intensities in the blood pool and remote myocardium to be able to account for intensity gradient proportional to the squared coil distance. Classification of myocardial intensities is performed using the EM-algorithm to overcome varying contrast and noise level between patients, centers and vendors. The EM-algorithm estimates the mean and standard deviation of intensity for normal myocardium and myocardium at risk based on the intensity histogram and was initialized based on the maximal extent model. Myocardium at risk was defined as a continuous region within the maximal perfusion territory of the culprit artery and assumed to be transmural. These *a priori* regional criteria were implemented by applying the classification by EM sector wise for sectors within the maximal extent model. The myocardium is divided into 24 sectors circumferentially. Further *a priori* information was implemented by using the infarct region from LGE images to define possible regions of microvascular obstruction as MaR despite the hypoenhancement. The original algorithm for T2-weighted images [12] was based on intensity classification by Expectation Maximization (EM) and utilization of a priori information on MaR. Surface coil intensity

correction and incorporation of the infarct region from LGE images was added in the new algorithm based on qualitative assessment of the CE-SSFP images. The new segmentation algorithm was named "Segment MaR CE-SSFP" and was implemented in the cardiac image analysis software Segment [9]. The algorithm will be made freely available at time of publication (<http://segment.heiberg.se>) and each processing block of the algorithm is further described in the Appendix.

## Comparison to other automatic threshold methods

The new automatic segmentation method was compared to three direct threshold methods which have been used for quantification of MaR in T2-weighted imaging [22-23], two standard deviations from remote (2SD) [24-25], full width half maximum intensity (FWHM) [26] and Otsu [27]. All methods used the same manual delineation of endocardium and epicardium. The 2SD threshold method estimates an intensity threshold from a remote region as the mean plus two standard deviations of the intensity within the remote region. The remote region was defined as the region outside the maximal extent model of the culprit artery [12]. The FWHM threshold method [10] estimates an intensity threshold from a remote region as midway between the mean intensity within the remote region and the maximal intensity within the myocardium. The remote region was defined in the same way as for 2SD. The threshold method of Otsu [28] estimates the intensity threshold from the histogram of all intensities to get minimal variance both above and below the threshold. For all three methods the intensity threshold was calculated and applied slice by slice as generally applied in T2-weighted images to account for the intensity gradient across slices.

## Statistical analysis

In the test set (n=183) quantification of MaR by the automatic Segment MaR CE-SSFP algorithm was compared to the manual delineation using Bland-Altman bias (mean  $\pm$  standard deviation), limits of agreement ([mean - 1.96 standard deviations; mean + 1.96 standard deviations]), and linear regression analysis (correlation coefficient). Regional agreement to manual delineation was evaluated by calculating Dice similarity coefficient (DSC) [29] (mean  $\pm$  standard deviation). Dice similarity coefficient can be derived from the kappa statistics for classification of pixels [30] and is calculated as two times the volume of the intersection of the MaR regions divided by the sum of the volumes of the MaR regions. The DSC is therefore 0 if the regions do not overlap and 1 if the regions overlap perfectly. Bias, linear regression and regional agreement was similarly analyzed for the three



automatic threshold methods, 2SD, FWHM and Otsu. Bias to manual delineation was analyzed separately for each of the three camera vendors for the automatic algorithm.

In a subset of 15 patients from the multicenter studies, inter-observer analysis of manual delineation and automatic segmentation was performed. Inter-observer analysis was assessed using Bland-Altman bias (mean  $\pm$  standard deviation), linear regression (correlation coefficient) and regional agreement DSC (mean  $\pm$  standard deviation) for manual delineation and automatic segmentation. Bias, linear regression and regional agreement was also assessed for automatic segmentation against manual delineation in the subset for comparison to inter-observer variability.

In the additional set (n=16), quantification of MaR in CE-SSFP images by the automatic Segment MaR CE-SSFP algorithm and manual delineation was compared to quantification of MaR in SPECT using bias (mean  $\pm$  standard deviation) and linear regression analysis (correlation coefficient).

The added value of each of the four processing blocks in the automatic algorithm described above was analyzed using bias (mean  $\pm$  standard deviation), linear regression analysis (correlation coefficient), regional agreement DSC (mean  $\pm$  standard deviation) and visualized by box-whisker plot of median, upper quartile, lower quartile, minimum, maximum and outliers. Two sided paired t-test of bias and DSC were performed for each processing block in comparison to the first block and the previous block with Bonferroni correction.

## Results

In the test set (n=183) MaR assessed by manual delineation in CE-SSFP was  $36 \pm 10$  % LVM and MaR assessed by Segment MaR CE-SSFP automatic segmentation was  $37 \pm 11$  %LVM. Bias was  $1 \pm 6$  %LVM [-11; 14] %LVM, R = 0.83 and regional agreement DSC  $0.85 \pm 0.08$  when Segment MaR CE-SSFP was compared to manual delineation (**Figure 2, Table 2**). **Figure 3** shows MaR in CE-SSFP at end-diastole and end-systole with manual delineation and automatic segmentation by Segment MaR CE-SSFP. The bias was lower, regression stronger and regional agreement higher for Segment MaR CE-SSFP than for the threshold methods of 2SD, FWHM and Otsu (**Figure 2, Table 2**). Bias to manual delineation analyzed per scanner vendors was  $0 \pm 7$  %LVM,  $2 \pm 6$  %LVM, and  $2 \pm 7$  %LVM, for automatic segmentation in images from GE (n=23), Philips (n=76), and Siemens (n=84), respectively. Inter-observer variability for manual delineation in CE-SSFP (n=15) was  $0 \pm 3$  %LVM compared to a bias

between manual delineation and Segment MaR CE-SSFP of  $2 \pm 6$  %LVM and inter-observer variability of Segment MaR CE-SSFP of  $-1 \pm 2$  %LVM (**Table 3**).

In the additional set of patients (n=16), MaR assessed by SPECT was  $27 \pm 10$  %LVM. In CE-SSFP MaR was by manual delineation  $28 \pm 7$  %LVM and by Segment MaR CE-SSFP  $29 \pm 7$  %LVM. Bias against SPECT was  $1 \pm 5$  %LVM (R=0.90) for CE-SSFP by manual reference delineation and  $2 \pm 7$  %LVM (R=0.73) by Segment MaR CE-SSFP (**Figure 4**).

**Table 2 - Results from test set n=183 for automatic Segment MaR CE-SSFP segmentation and threshold methods against manual delineation**

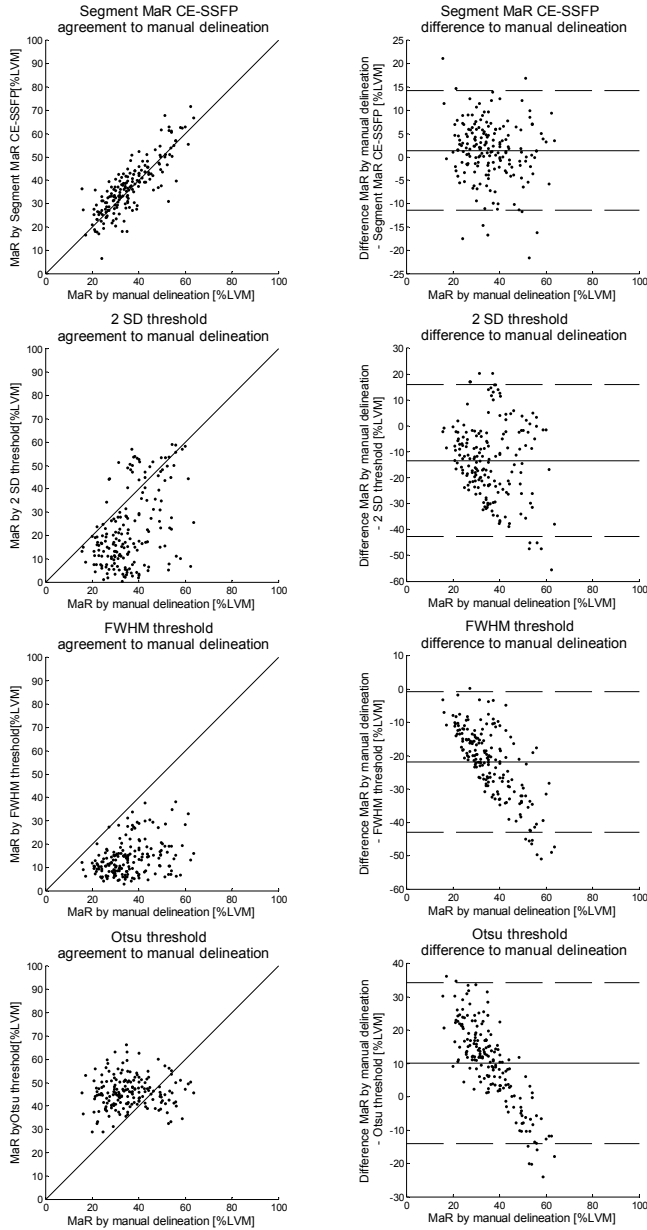
	Bias %LVM	R- value	DSC
Segment MaR CE-SSFP	$1 \pm 6$	0.83	$0.85 \pm 0.08$
2SD threshold	$-13 \pm 15$	0.47	$0.54 \pm 0.27$
FWHM threshold	$-22 \pm 11$	0.42	$0.42 \pm 0.21$
Otsu threshold	$10 \pm 12$	0.05	$0.65 \pm 0.12$

MaR-Myocardium at risk, LVM- Left ventricular mass, DSC- Dice similarity coefficient, Segment MaR CE-SSFP - automatic segmentation proposed in this study, 2SD-two standard deviations from remote, FWHM- full width half maximum intensity

**Table 3 - Inter-observer variability analysis from subset n=15 for manual delineation and automatic Segment MaR CE-SSFP segmentation compared to results for Segment MaR CE-SSFP against manual delineation**

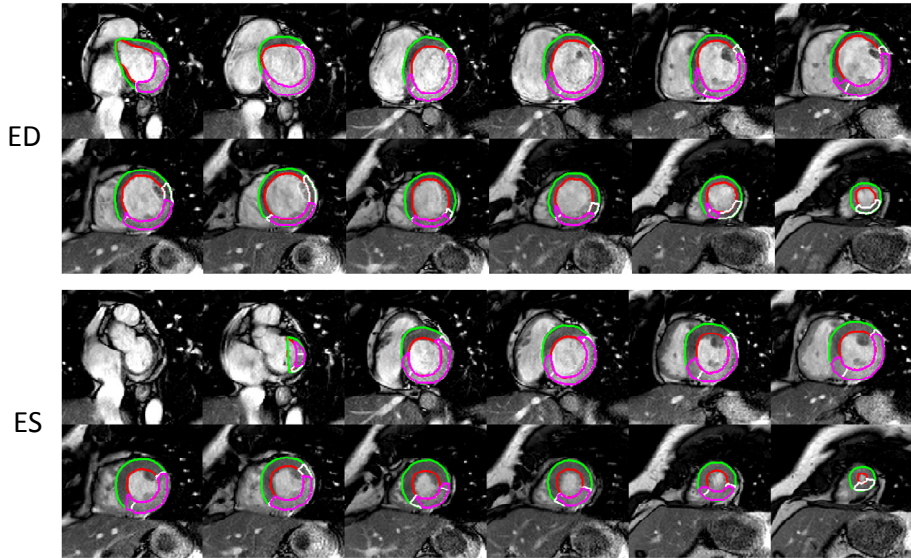
	Bias % LVM	R- value	DSC
Manual delineation vs. manual delineation	$0 \pm 3$	0.93	$0.92 \pm 0.04$
Segment MaR CE-SSFP vs. Segment MaR CE-SSFP	$-1 \pm 2$	0.99	$0.94 \pm 0.03$
Segment MaR CE-SSFP vs. manual delineation	$2 \pm 6$	0.77	$0.86 \pm 0.05$

MaR-Myocardium at risk, LVM- Left ventricular mass, DSC- Dice similarity coefficient, Segment MaR CE-SSFP - automatic segmentation proposed in this study, manual delineation performed by a reference and a second observer, automatic Segment MaR CE-SSFP performed by a reference and a second observer



**Figure 2 - Correlation and bias for automatic segmentation and threshold methods against manual delineation**

Correlation of MaR as % of LVM (left column) and Bland-Altman plot of MaR bias as % of LVM (right column) for the automatic segmentation algorithm (first row), threshold of 2SD from remote (second row), FWHM (third row) and Otsu (fourth row), all against manual delineation. The line of identity is shown as a solid line for all correlations plots and mean bias (solid line) and mean  $\pm$  two standard deviations (dashed line) is shown for all Bland-Altman plots.



**Figure 3 - Example of automatic segmentation and manual delineation of MaR in CE-SSFP**

Typical MaR segmentation in all left ventricular short axis slice images from one patient in end-diastole (ED, top panel) and end systole (ES, bottom panel), for automatic segmentation by Segment MaR CE-SSFP, shown in white, and manual delineation, shown in purple. Endocardial borders are shown in red and epicardial border in green. For this patient MaR by manual segmentation was 44 %LVM and by automatic Segment MaR CE-SSFP 43% LVM with a regional agreement DSC of 0.85.

A significant difference in regional agreement DSC was shown for each of the processing blocks of the Segment MaR CE-SSFP algorithm even though the difference in bias %LVM was not significant (Figure 5, Table 4).

## Discussion

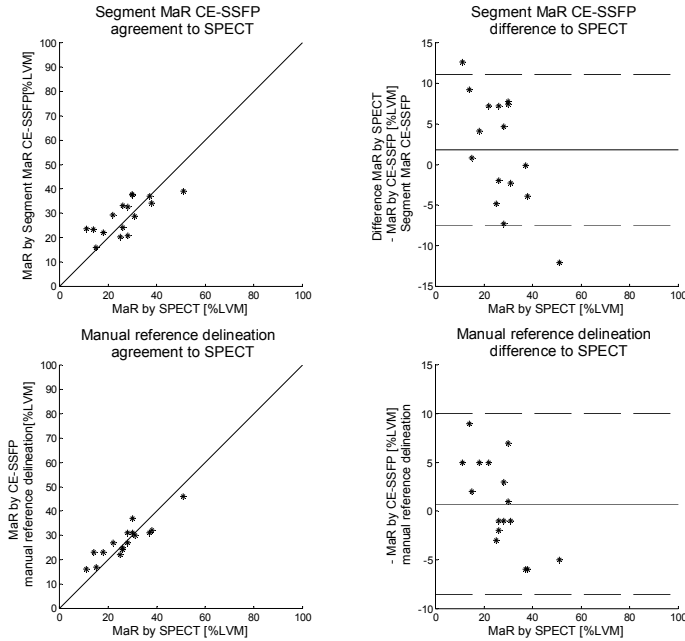
This study has presented an automatic algorithm for quantification of MaR in CE-SSFP images, validated against manual delineation in 183 patients from two multi-center, multi-vendor studies and against SPECT, as reference method, in 16 patients. The proposed automatic segmentation, Segment MaR CE-SSFP, shows low bias and variability, strong correlation and high regional agreement compared to manual delineation and SPECT. The Segment MaR CE-SSFP algorithm was shown superior to thresholding methods (2SD, FWHM and Otsu).

## Technical aspects

The added value of each processing block was shown significant by regional agreement DSC even though a significant difference in bias was only seen when bias changed from overestimation to underestimation adding use of a *priori* information

on MaR. The significant change seen for DSC highlights the importance of analyzing regional agreement as a part of the validation in addition to bias.

Expectation Maximization was shown superior to 2SD, FWHM and Otsu, when considering regional agreement DSC, quantitative bias and correlation R-value. The EM-algorithm was used by Johnstone et al. [13] to find the mean and standard deviation of remote myocardium in T2-weighted black blood images, but the mean and standard deviation of edema was not used to define the threshold which may explain the lower regional agreement with DSC  $0.50 \pm 0.27$ . Gao et al. [31] also used the EM-algorithm to classify myocardial intensities in T2-weighted bright blood images, with the assumption of Rayleigh-Gaussian mixture model. Rayleigh distributed intensities were assumed due to nulling of remote myocardium [31] which is not done in CE-SSFP and therefore, in this study, Gaussian intensity distributions were assumed for both normal myocardium and MaR. Surface coil intensity correction was shown to increase regional agreement. The surface coil correction was based on intensities in remote myocardium and blood pool and thereby the bright blood property of CE-



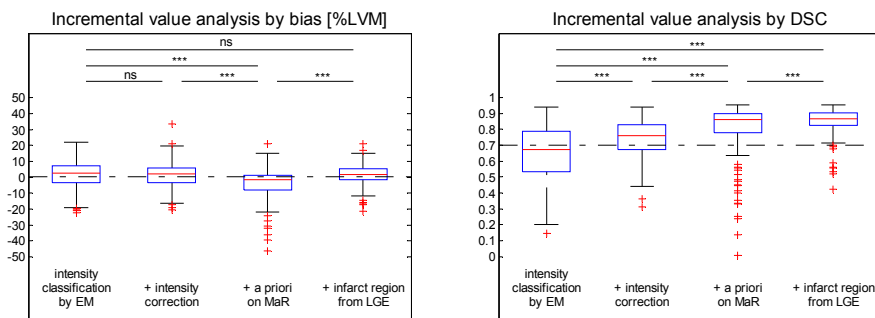
**Figure 4 Correlation and bias against SPECT for automatic segmentation and manual delineation in CE-SSFP**

Correlation of MaR as % of LVM (left column) and Bland-Altman plot of MaR bias as % of LVM (right column) against SPECT for automatic segmentation algorithm Segment MaR CE-SSFP (top row) and manual reference delineation (bottom row). The line of identity is shown as a solid line for all correlations plots and mean bias (solid line) and mean  $\pm$  two standard deviations (dashed line) is shown for all Bland-Altman plots. Correlation and Bland-Altman plots for manual delineation in CE-SSFP against SPECT (bottom row) are adopted from Sorenson et al. [4].

SSFP was advantageous to the black blood T2-STIR images in the original Segment MaR algorithm [12] where no intensity correction was applied. Gao et al. [31] used intensity correction developed for the bright blood ACUT2E [32] images with use of proton density maps and achieved a DSC  $0.7 \pm 0.06$  before applying feature analysis. By incorporating *a priori* regional criteria in the definition of the MaR region the regional agreement was further increased from 0.74 to 0.81. This is in line with Gao et al. [31] who showed increased regional agreement by DSC from 0.7 to 0.74 by adding feature analysis of the edema region. Both regional agreement by DSC and quantitative bias as %LVM was improved by the addition of information on the infarct region from LGE images which has not been implemented in previous studies. As for all automatic segmentation methods visual assessment and possibly manual corrections are needed and will probably influence the outliers seen after the fourth block of the algorithm and decrease the variability further.

### Comparison to previous studies

Regional agreement to manual delineation was for Segment MaR CE-SSFP higher than for the automatic segmentation methods by Johnstone et al. [13] (DSC  $0.50 \pm 0.27$ ) and Gao et al. [31] (DSC  $0.74 \pm 0.06$ ), and higher respectively similar to inter-observer regional agreement of manual delineation in the same studies (DSC  $0.72 \pm 0.14$  [13] and  $0.85 \pm 0.03$  [31]). Regional agreement of interobserver variability bias of Segment MaR CE-SSFP was comparable to inter-observer variability of manual delineation found in this study and similar to inter-observer variability previously found in CE-SSFP ( $2 \pm 4$  %LVM [4] and  $0 \pm 6$  %LVM [18]), and in T2-weighted imaging ( $-2 \pm 5$  %LVM [12] and  $5 \pm 5$  %LVM [18]). Bias of Segment MaR CE-SSFP to SPECT was low and comparable to the results from the validation study of CE-SSFP with manual delineation ( $0 \pm 5$  %LVM [4]). Bias was lower and regional agreement was higher for Segment MaR CE-SSFP than for the threshold methods of 2SD, FWHM and Otsu.



**Figure 5 - Analysis of incremental value of blocks in the automatic segmentation algorithm**

Incremental value of each block in the automatic segmentation algorithm analyzed by bias to manual delineation as %LVM, left panel and by regional agreement as Dice similarity coefficient DSC(right panel). Bias and DSC was calculated with segmentation based on only intensity classification by Expectation Maximization and calculated after the addition of the processing blocks of intensity correction, a priori on myocardium at risk (MaR) and infarct region from late gadolinium enhancement (LGE). For each block of the algorithm the upper limit of the box indicate upper quartile, middle line indicate median, lower limit of box indicate lower quartile, whiskers indicate minimum and maximum and points (+) indicate outliers. Bias zero is shown as dotted black line in the left panel, DSC above of 0.7 indicates good regional agreement [30], and is shown as dotted black line in the right panel. Two sided paired t-test was performed for each block in comparison to previous block and first block, ns: non significant, \*\*\*:  $p < 0.0001$ .

**Table 4 - Analysis of incremental value of each block in the automatic Segment MaR CE-SSFP algorithm (n=183)**

	Bias %LVM	R- value	DSC
Intensity classification by EM	$2 \pm 8$	0.60	$0.65 \pm 0.18$
+ intensity correction	$2 \pm 8$	0.63	$0.74 \pm 0.12$
+ a priori on MaR	$-4 \pm 10$	0.62	$0.81 \pm 0.16$
+ infarct region from LGE	$1 \pm 6$	0.83	$0.85 \pm 0.08$

EM- Expectation Maximization, MaR-Myocardium at risk, LGE- late gadolinium enhancement, LVM- Left ventricular mass, DSC- Dice similarity coefficient

Recently McAlindon et al. [22] showed that manual delineation in T2-weighted images was superior to simple threshold methods (2, 3 and 5 SD from remote, FWHM and Otsu) with manual corrections regarding accuracy and variability of intra-observer, inter-observer and test-retest. Khan et al. [23] also showed that using simple threshold methods with manual corrections for inclusion of hypoenhancement and exclusion of artifacts did not significantly reduce time for analysis compared to

manual delineation. Automatic segmentation methods are desirable to increase accuracy, decrease subjectivity and reduce time for analysis. Using simple threshold methods for MaR quantification in T2-weighted images seems to achieve neither and might be explained by the regional agreement seen in T2-weighted images (DSC  $0.69 \pm 0.14$ ,  $0.46 \pm 0.14$  and  $0.68 \pm 0.10$  for 2SD, FWHM and Otsu respectively) [12] and in CE-SSFP images of this study the regional agreement for simple threshold methods was even lower. The regional agreement of the Segment MaR CE-SSFP was however similar to that of the original Segment MaR in T2-weighted images (DSC  $0.85 \pm 0.07$ ) [12].

Segment MaR CE-SSFP was designed to include hypoenhancement and exclude artifacts and thereby has a greater potential to reduce time for analysis, and with a low bias to manual delineation by expert readers and a regional agreement and bias to manual delineation comparable to inter-observer of manual delineation Segment MaR CE-SSFP shows potential to increase accuracy and reduce subjectivity.

## Limitations

Limitations to the study are that test-retest scans were not performed and the effect of and time required for possible manual corrections following automatic segmentation was not evaluated. Contrast enhanced SSFP are not yet widely used for assessment of MaR but has been shown to be more robust than T2-weighted imaging in multi-center,

multi-vendor studies [8] and can easily be implemented by acquiring cine SSFP images approximately 5 minutes after gadolinium injection.

## Conclusion

This study has presented an automatic algorithm, Segment MaR CE-SSFP for quantification of MaR in CE-SSFP images based on four processing blocks, Expectation Maximization, surface coil intensity correction, *a priori* regional criteria and incorporation of infarct region from LGE images. Low bias and variability, strong correlation and high regional agreement was shown against manual delineation in CE-SSFP images from multi-center, multi-vendor randomized clinical trials. Bias and variability was comparable to inter-observer variability of manual delineation and inter-observer variability was decreased by use of the Segment MaR CE-SSFP algorithm.

## Abbreviations

CE-SSFP : contrast enhanced SSFP, DSC : Dice similarity coefficient, EM : expectation maximization, FWHM : full width half maximum, LAD : left anterior descending artery, LCx : left circumflex artery, LGE : late gadolinium enhancement, LM : left main artery, MaR : myocardium at risk, MI : myocardial infarction, MSI : myocardial salvage index, RCA : right coronary artery, RCT : randomized clinical trials, SPECT : Single photon emission computed tomography, %LVM : percent of left ventricular mass

## Acknowledgements

This study has been funded by the Swedish Research Council (2011-3916, 2012-4944, 2011-4078) , The Swedish Heart and Lung Foundation, The Medical Faculty of Lund University, Sweden, and Region of Scania, Sweden. Funding for the CHILL-MI trial was received from Philips Healthcare. Funding for the MITOCARE trial was received from the European Commission within the 7<sup>th</sup> Framework Programme for RTD – Project MITOCARE – Grant Agreement HEALTH-2010-261034.

## Author details

<sup>1</sup>Dept. of Clinical Physiology, Skåne University Hospital in Lund, Lund University, Lund, Sweden,

<sup>2</sup>Dept. of Biomedical Engineering, Faculty of Engineering, Lund University, Lund, Sweden,

<sup>3</sup>Laboratory of Medical Informatics, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki , Greece, <sup>4</sup>Dept. of Cardiology, Henri Mondor Hospital, Creteil, France, <sup>5</sup>Dept. of Cardiology, Lund University, Lund, Sweden, <sup>6</sup>Dept. of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden, <sup>7</sup>Dept. of

Cardiology B, Oslo, University Hospital Ullevål and Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding author: Einar Heiberg Department of Clinical Physiology, Skåne University Hospital, Lund, SE 221 85 Lund

## Authors' contributions

JT contributed to the design of the study, developed and implemented the automatic segmentation algorithm, analyzed and interpreted results, and drafted the manuscript. MC and HE contributed to the design of the study and the automatic algorithm and performed manual delineations in the test set. AA contributed to the design of the study and the automatic algorithm and provided in depth CMR knowledge needed for algorithm development. JFD, SK, DA and DE were responsible for data collection in the multi-center studies. PS was responsible for data collection for validation against SPECT and performed manual delineations in the additional set. JP was responsible for data collection for validation against SPECT. HA contributed to the design of the automatic algorithm, performed manual delineations and conceived the study. EH contributed to the design of the automatic algorithm and conceived the study. All authors revised the manuscript, and have read and approved the final version of the manuscript.

## Competing interests

EH is the founder of Medviso AB, Lund, Sweden, which sells a commercial version of Segment. JT is employed by Medviso AB on a part-time basis. HA is a share-holder of Imacor AB, Lund, Sweden, which performs core lab analysis of CMR images. HA, HE and MC have been employed by Imacor AB on a part-time basis. The other authors declare that they have no competing interests.

## Additional files

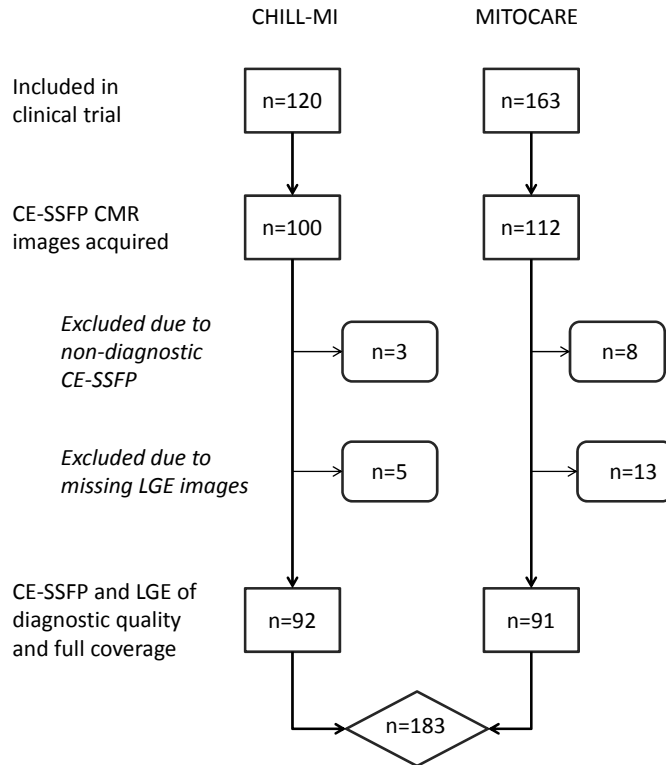
Supplemental figure 1- Patient inclusion from clinical trials

Appendix 1 - Detailed description of the automatic segmentation method

## References:

1. Engblom, H., et al., Design of clinical cardioprotection trials using CMR: impact of myocardial salvage index and a narrow inclusion window on sample size. *Journal of Cardiovascular Magnetic Resonance*, 2015. 17(Suppl 1): p. P90.
2. Kim, R.J., et al., Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. *Circulation*, 1999. 100(19): p. 1992-2002.
3. Aletras, A.H., et al., Retrospective determination of the area at risk for reperfused acute myocardial infarction with T2-weighted cardiac magnetic resonance imaging: histopathological and displacement encoding with stimulated echoes (DENSE) functional validations. *Circulation*, 2006. 113(15): p. 1865-70.

4. Sorensson, P., et al., Assessment of myocardium at risk with contrast enhanced steady-state free precession cine cardiovascular magnetic resonance compared to single-photon emission computed tomography. *J Cardiovasc Magn Reson*, 2010. 12(1): p. 25.
5. Carlsson, M., et al., Myocardium at risk after acute infarction in humans on cardiac magnetic resonance: quantitative assessment during follow-up and validation with single-photon emission computed tomography. *JACC Cardiovasc Imaging*, 2009. 2(5): p. 569-76.
6. Erlinge, D., et al., Rapid endovascular catheter core cooling combined with cold saline as an adjunct to percutaneous coronary intervention for the treatment of acute myocardial infarction. The CHILL-MI trial: a randomized controlled study of the use of central venous catheter core cooling combined with cold saline as an adjunct to percutaneous coronary intervention for the treatment of acute myocardial infarction. *J Am Coll Cardiol*, 2014. 63(18): p. 1857-65.
7. Atar, D., et al., Effect of intravenous TRO40303 as an adjunct to primary percutaneous coronary intervention for acute ST-elevation myocardial infarction: MITOCARE study results. *Eur Heart J*, 2015. 36(2): p. 112-9.
8. Nordlund, D., et al., Performance of contrast enhanced SSFP and T2-weighted imaging for determining myocardium at risk in a multi-vendor, multi-center setting- data from the MITOCARE and CHILL-MI trials. *Journal of Cardiovascular Magnetic Resonance*, 2015. 17(Suppl 1): p. P194.
9. Heiberg, E., et al., Automated quantification of myocardial infarction from MR images by accounting for partial volume effects: animal, phantom, and human study. *Radiology*, 2008. 246(2): p. 581-8.
10. Hsu, L.Y., et al., Quantitative myocardial infarction on delayed enhancement MRI. Part I: Animal validation of an automated feature analysis and combined thresholding infarct sizing algorithm. *J Magn Reson Imaging*, 2006. 23(3): p. 298-308.
11. Amado, L.C., et al., Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. *J Am Coll Cardiol*, 2004. 44(12): p. 2383-9.
12. Sjogren, J., et al., Semi-automatic segmentation of myocardium at risk in T2-weighted cardiovascular magnetic resonance. *J Cardiovasc Magn Reson*, 2012. 14: p. 10.
13. Johnstone, R.I., et al., Assessment of tissue edema in patients with acute myocardial infarction by computer-assisted quantification of triple inversion recovery prepared MRI of the myocardium. *Magn Reson Med*, 2011.
14. Rationale and design of the 'MITOCARE' Study: a phase II, multicenter, randomized, double-blind, placebo-controlled study to assess the safety and efficacy of TRO40303 for the reduction of reperfusion injury in patients undergoing percutaneous coronary intervention for acute myocardial infarction. *Cardiology*, 2012. 123(4): p. 201-7.
15. Simonetti, O.P., et al., An improved MR imaging technique for the visualization of myocardial infarction. *Radiology*, 2001. 218(1): p. 215-23.
16. Heiberg, E., et al., Design and validation of Segment--freely available software for cardiovascular image analysis. *BMC Med Imaging*, 2010. 10: p. 1.
17. Beek, A.M., R. Nijveldt, and A.C. van Rossum, Intramyocardial hemorrhage and microvascular obstruction after primary percutaneous coronary intervention. *Int J Cardiovasc Imaging*, 2010. 26(1): p. 49-55.
18. Ubachs, J.F., et al., Myocardium at risk by magnetic resonance imaging: head-to-head comparison of T2-weighted imaging and contrast-enhanced steady-state free precession. *Eur Heart J Cardiovasc Imaging*, 2012. 13(12): p. 1008-15.
19. Ugander, M., et al., Quantification of myocardium at risk in myocardial perfusion SPECT by co-registration and fusion with delayed contrast-enhanced magnetic resonance imaging--an experimental ex vivo study. *Clin Physiol Funct Imaging*, 2012. 32(1): p. 33-8.
20. Soneson, H., et al., An improved method for automatic segmentation of the left ventricle in myocardial perfusion SPECT. *J Nucl Med*, 2009. 50(2): p. 205-13.
21. Dempster, A.P., N.M. Laird, and D.B. Rubin, Maximum Likelihood from Incomplete Data Via Em Algorithm. *Journal of the Royal Statistical Society Series B-Methodological*, 1977. 39(1): p. 1-38.
22. McAlindon, E., et al., Quantification of infarct size and myocardium at risk: evaluation of different techniques and its implications. *Eur Heart J Cardiovasc Imaging*, 2015.
23. Khan, J.N., et al., Comparison of semi-automated methods to quantify infarct size and area at risk for cardiovascular magnetic resonance imaging at 1.5T and 3.0T field strengths. *BMC Res Notes*, 2015. 8(1): p. 1007.
24. Friedrich, M.G., et al., The salvaged area at risk in reperfused acute myocardial infarction as visualized by cardiovascular magnetic resonance. *J Am Coll Cardiol*, 2008. 51(16): p. 1581-7.
25. Wright, J., et al., Quantification of myocardial area at risk with T2-weighted CMR: comparison with contrast-enhanced CMR and coronary angiography. *JACC Cardiovasc Imaging*, 2009. 2(7): p. 825-31.
26. Tilak, G.S., et al., In vivo T2-weighted magnetic resonance imaging can accurately determine the ischemic area at risk for 2-day-old nonreperfused myocardial infarction. *Invest Radiol*, 2008. 43(1): p. 7-15.
27. Burchell, T., et al., Comparing analysis methods for quantification of myocardial oedema in patients following reperfused ST-elevation MI. *Journal of Cardiovascular Magnetic Resonance*, 2011. 13(Suppl 1): p. M11.
28. Otsu, N., Threshold Selection Method from Gray-Level Histograms. *Ieee Transactions on Systems Man and Cybernetics*, 1979. 9(1): p. 62-66.
29. Dice, L.R., Measures of the Amount of Ecologic Association between Species. *Ecology*, 1945. 26(3): p. 297-302.
30. Zijdenbos, A.P., et al., Morphometric Analysis of White-Matter Lesions in Mr-Images - Method and Validation. *Ieee Transactions on Medical Imaging*, 1994. 13(4): p. 716-724.
31. Gao, H., et al., Highly automatic quantification of myocardial oedema in patients with acute myocardial infarction using bright blood T2-weighted CMR. *J Cardiovasc Magn Reson*, 2013. 15: p. 28.
32. Aletras, A.H., et al., ACUT2E TSE-SSFP: a hybrid method for T2-weighted imaging of edema in the heart. *Magn Reson Med*, 2008. 59(2): p. 229-35.



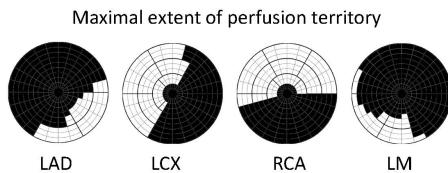
**Supplemental figure 1- Patient inclusion from clinical trials**

Patient inclusion from clinical trials CHILL-MI and MITOCARE resulted in 183 patients in the test set. In total 29 patients with CE-SSFP images were excluded due to non-diagnostic image quality or missing LGE images.



## Appendix

The automatic segmentation Segment MaR CE-SSFP was developed for segmentation of MaR in CE-SSFP based on ideas from the algorithm developed for T2-weighted images [12]. The use of Expectation Maximization (EM) algorithm [21] for classification of myocardial intensities [12] was improved with modified constraints and surface coil intensity correction. For definition of the MaR region the implementation of *a priori* regional criteria [12] was improved and utilization of information on infarct region from LGE images was added. Maximal extent models for the perfusion territory of the culprit artery [12] were used to define the remote region and the culprit region based on user input. **Figure A1** shows the maximal extent models as defined in consensus by three experienced observers [12]. As input to the automatic algorithm, the user defines the culprit artery as either LAD, LCx, RCA or LM based on the overall appearance of the hyper enhanced region and defines right ventricular insertion points in CE-SSFP and LGE images, to define maximal extent model and how to rotate the model. The maximal extent model is used with the user input of culprit artery and LV rotation to define the remote myocardium for surface coil intensity correction and initialization of the EM-algorithm and to define a MaR region within the maximal perfusion territory.

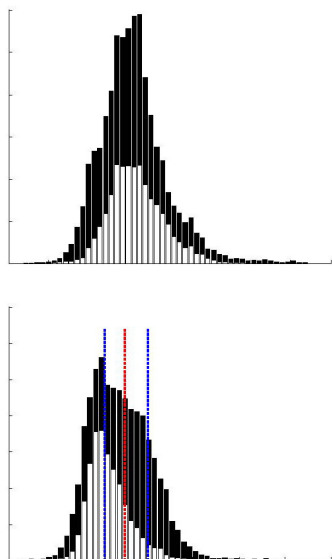


**Figure A1 - Model of maximal extent for perfusion territory of each culprit artery**

*Bulls-eye representation of maximal extent model for the perfusion territories of left anterior descending artery (LAD), left circumflex artery (LCx), right coronary artery (RCA), and left main artery (LM). Models for LAD, LCx and RCA were defined in consensus by three experienced observers in an extended 17-segment AHA model and models for LM were defined from the models of LAD, LCx and RCA. The 17-segment model is extended to three slices in each of the basal, mid-ventricular and apical zones and 24 sectors in each slice. Black sectors are included in the maximal extent model. The septal part of the left ventricle is represented in the left of the bulls-eye plot, the lateral part in the right, anterior part in the top, inferior part in the bottom, the apical slices in the center and the basal slices in the outer part of the bulls-eye plot.*

Varying surface coil sensitivity may cause an intensity gradient through the CMR images and can in CE-SSFP cause a larger variability in the myocardium than the contrast between MaR and normal myocardium and hence a surface coil correction needs to be applied before the EM-algorithm. A second order intensity correction is applied to account for a gradient proportional to the squared distance to the surface coil. The correction is calculated based on the intensity in the remote myocardium and blood pool with papillaries excluded from the blood pool by using a simple unconstrained EM-algorithm. The intensity correction should result in a reduced intensity variability in the remote myocardium and a mean intensity in the culprit region higher than in the remote region, otherwise the correction is not applied. If the mean intensity in the remote myocardium is higher than in the culprit region both before and after the intensity correction, no correction is applied and the user is notified with a warning on low image quality. **Figure A2** shows the intensity histogram before and after intensity correction for the remote and culprit region.

For classification of pixel intensities as normal myocardium or MaR, a Bayesian probability is calculated by the use of a constrained EM-algorithm. The EM-algorithm [21] iteratively refines an initial classification to find the maximum likelihood estimate of the mean and standard deviation for the intensity distributions of normal myocardium and MaR. The initial classification is defined from the maximal extent model with all pixels in the remote region initially classified as normal myocardium and all pixels in the culprit region initially classified as MaR. The EM-algorithm was constrained to keep the initial classification of normal myocardium for pixels with intensity below the 50th percentile in the remote region, respectively, keeping classification of MaR for pixels with intensity above the 75th percentile in the culprit region. The Bayesian MaR probability is calculated for each myocardial pixel as the intensity distribution of MaR divided by the sum of the intensity distributions of MaR and normal myocardium. The resulting Bayesian MaR probability cutoff 0.5 indicates higher probability of MaR and is shown in the histogram after intensity correction in **Figure A2**.



**Figure A2 - Surface coil intensity correction**

Histogram of myocardial intensities within the myocardium before intensity correction (top panel) and after intensity correction (bottom panel), represented in black for culprit region and white for remote region. After intensity correction a decreased standard deviation of remote region is obtained and the mean intensity of the culprit region is higher than remote region. Myocardial intensities as after intensity correction (bottom panel) are used as input to the EM-algorithm with remote region and culprit region as initial classification and dashed blue lines indicating the constraints. Intensities below the 50th percentile of the remote region are kept classified as normal myocardium through the iterations of the EM-algorithm (lower dashed blue line). Intensities above the 75th percentile of the intensities within the culprit region are kept classified as MaR through the iterations of the EM-algorithm (upper dashed blue line). Dashed red line indicates the resulting cut off 0.5 of the Bayesian probability of MaR resulting from the EM-algorithm after intensity correction.

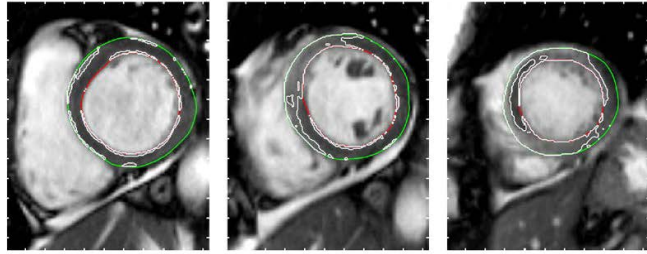
The MaR region is defined as a connected region with high MaR probability which fulfills the *a priori* criteria on transmural and localization within the culprit artery's perfusion territory. The mean MaR probability is calculated for each sector in a bullseye representation with 24 sectors and 30 interpolated slices averaged over the time frames.

Sectors with a mean MaR probability above 0.5 which are within the maximal extent and connected to its nearest neighboring sector within the slice or in an adjacent slice in a 4-neighbourhood constitute a region. If several connected regions are found the region with highest summed MaR probability is chosen. Gray scale morphological operations of opening and closing are applied to remove holes and small peninsulas in a 4-neighbourhood. Additionally holes within slices are removed to account for larger encapsulated regions of microvascular obstruction. Non-physiological extent in apical and basal slices is detected for LAD and LM as missing apical MaR sectors or false basal sectors and for LCx and RCA as false apical sectors. False and missing sectors was detected as extent larger than mean + 2 standard deviations respectively smaller than mean - 2 standard deviations of the extent in midventricular slices. If any non-physiological extent was detected and corrected for, then the user is notified to check correctness of MaR region in the basal or apical slices. Figure A3 shows one short axis slice with MaR segmentation before and after applying *a priori* regional criteria.

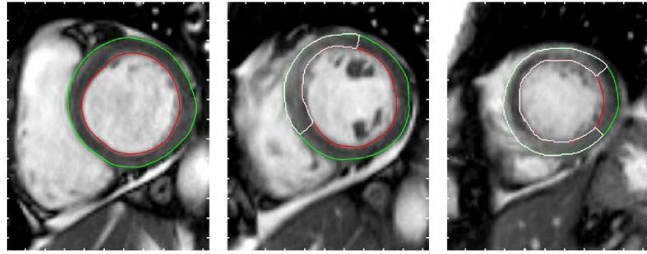
If LGE images with infarct segmentation are available the information on the infarct region can be used as part of the *a priori* regional criteria for definition of the MaR region. The infarct region is always a part of MaR but may due to hypoenhancement of microvascular obstruction not always be detected as MaR by the EM-algorithm. From the LGE images with delineation of the infarct, either by manual delineation or automatic segmentation [9], and right ventricular insertion points the infarct region, represented as a sector-wise bullseye, is used to define the MaR region. For each sector the fraction of infarct is calculated and the MaR region is defined from sectors with either the infarct fraction above 0.5 or mean MaR probability above 0.5. Figure A4 shows a short axis slice of a CE-SSFP image with a distinct region of microvascular obstruction which can be determined as MaR region by the use of the infarct segmentation from the LGE images.

From the bulls eye representation of the MaR region a MaR segmentation is defined in the short axis slices for each time frame and MaR is expressed as %LVM averaged over end-diastole and end-systole.

Without *a priori*  
regional criteria



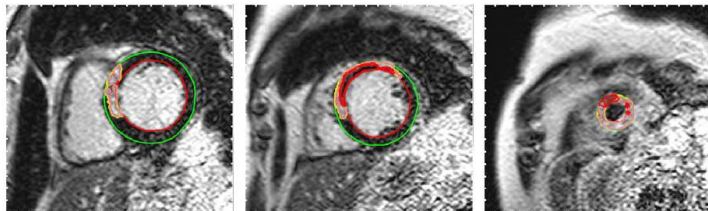
With *a priori*  
regional criteria



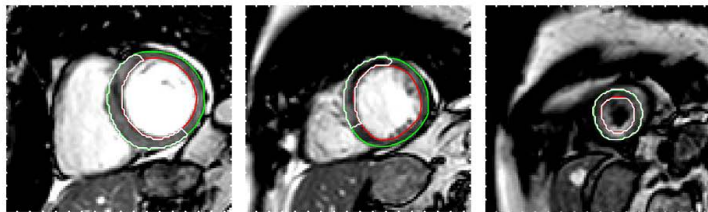
**Figure A3 - Utilization of *a priori* information on extent and transmuralty**

Three short axis slice, basal, mid and apical with MaR segmentation, shown in white, based on only the Bayesian probability of MaR (top row) and with the addition of *a priori* regional criteria on extent and transmuralty (bottom row).

Infarct  
by LGE



MaR  
by CE-SSFP



**Figure A4 - Utilization of infarct segmentation from LGE**

Three short axis slices, basal, midventricular and apical, from LGE (top row) with infarct region in yellow, infarct core in pink and microvascular obstruction in red, and CE-SSFP (bottom row) with automatic segmentation by Segment MaR CE-SSFP after utilization of information on infarct region from LGE. The large region of hypoenhancement in CE-SSFP is included as MaR by utilizing the segmentation of infarct region.



## Paper V



# A new automatic algorithm for quantification of myocardial infarction in late gadolinium enhancement cardiac magnetic resonance images: experimental validation and implementation in multi-center, multi-vendor patient data

Jane Tufvesson<sup>1,2</sup>, Robert Jablonowski<sup>1</sup>, Henrik Engblom<sup>1</sup>, Marcus Carlsson<sup>1</sup>, Anthony H. Aletras<sup>1,3</sup>, Pavel Hoffmann<sup>4</sup>, Alexis Jacquier<sup>5</sup>, Frank Kober<sup>6</sup>, Bernhard Metzler<sup>7</sup>, David Erlinge<sup>8</sup>, Dan Atar<sup>4</sup>, Håkan Arheden<sup>1</sup>, Einar Heiberg<sup>1,2</sup>

## Abstract

### Background

Late gadolinium enhancement (LGE) cardiovascular magnetic resonance (CMR) using magnitude inversion recovery (IR) or phase sensitive inversion recovery (PSIR) has become clinical standard for assessment of myocardial infarction (MI). However, there is no clinical standard for quantification of MI even though multiple methods have been proposed. Simple thresholds have yielded varying results and advanced algorithms have only been validated in single center studies. Therefore, the aim of this study was to develop an automatic algorithm for MI quantification in IR and PSIR LGE images and to validate the new algorithm experimentally and apply the new algorithm in multi-center, multi-vendor patient data.

### Methods

The new automatic algorithm, EWA (Expectation Maximization, weighted intensity, a priori information), was implemented using an intensity threshold by Expectation Maximization (EM) and a weighted summation to account for partial volume effects.

The EWA algorithm was validated *in-vivo* against triphenyltetrazolium-chloride (TTC) staining (n=6 pigs with paired IR and PSIR images) and against *ex-vivo* high resolution T1-weighted images (n=20 IR and n=12 PSIR images). The EWA algorithm was compared to expert delineation in 124 patients from multi-center, multi-vendor clinical trials 2-6 days following first time ST-elevation myocardial infarction (STEMI) treated with percutaneous coronary intervention (PCI) (n=75 IR and n=49 PSIR images).

### Results

Infarct size by the EWA algorithm showed a bias to TTC of  $-1 \pm 1\%$ LVM (R=0.99) in IR and  $-2 \pm 3\%$ LVM (R=0.88) in PSIR images and a bias to *ex-vivo* T1-weighted images of  $-1 \pm 3\%$ LVM (R=0.96) in IR and  $0 \pm 5\%$ LVM (R=0.77) in PSIR images. In multi-center patient studies, infarct size by the EWA algorithm was  $15 \pm 8\%$ LVM in IR images (n=124) and  $17 \pm 10\%$ LVM in PSIR images (n=49) with a bias to expert delineation of  $-2 \pm 6\%$ LVM (R=0.81) and  $-1 \pm 5\%$ LVM (R=0.89), respectively.

### Conclusions

The EWA algorithm was validated experimentally and applied in patient data with a low bias in both IR and PSIR LGE images. Thus, the use of EM and a weighted intensity as in the EWA algorithm, may serve as a clinical standard for the quantification of myocardial infarction in LGE CMR images.

## Background

Late gadolinium enhancement (LGE) cardiovascular magnetic resonance (CMR) is considered the reference standard for the assessment of myocardial infarction (MI) [1-2]. Visualization of MI by use of gadolinium

enhancement has evolved from T1-weighted imaging in 1984 [3] to current use of LGE magnitude inversion recovery (IR) [4] and phase sensitive inversion recovery (PSIR) sequences [5] as clinical standard [6].

Although multiple methods have been proposed for quantification of MI, there is still no clinical standard [6]. Manual delineation or visual grading of MI is often used clinically but has the disadvantage of being subjective and therefore threshold techniques have been proposed based on different numbers of standard deviations (SD) from remote myocardium or based on the full width half maximum (FWHM) intensity threshold [7-9]. Recently Stirrat *et al.* [10] showed a difference between infarct size derived from IR and PSIR LGE images for threshold methods of SD from remote and FWHM. More advanced methods for MI quantification has been implemented and validated as the FACT algorithm by Hsu *et al.* [11-12] and the weighted algorithm by Heiberg *et al.* [13]. Both algorithms involve regional analysis of the infarcted myocardium to include microvascular obstruction (MVO) and exclude artifacts. However, the FACT algorithm [11] was developed and validated for PSIR images with surface coil intensity correction and based on a FWHM threshold, whereas the weighted algorithm [13] was developed and validated for magnitude IR images. Heiberg *et al.* [13] used a weighted approach to account for partial volume effects which was shown to decrease variability compared to the use of pure signal intensity thresholds. The algorithm was, however, based on a SD threshold from remote and the weighted approach was not applied in *ex-vivo* high resolution T1-weighted images. Using a threshold by Expectation Maximization (EM) [14] has been shown superior to FWHM and SD from remote for quantification of myocardium at risk in T2-weighted images [15] and the EM-algorithm has also been implemented for MI quantification in LGE images [16]. The EM-algorithm has previously not been combined with a weighted approach and, to the best of our knowledge, no algorithms have been developed for MI quantification in both IR and PSIR LGE images and applied in multi-center, multi-vendor patient studies.

Therefore, the aim of this study was 1) to develop a new automatic algorithm for MI quantification by combining intensity threshold by Expectation Maximization (EM) with a weighted approach to account for partial volume effects, 2) to validate the automatic algorithm experimentally for IR and PSIR LGE images and *ex-vivo* high resolution T1-weighted images, and 3) apply the automatic algorithm in multi-center, multi-vendor patient data with consensus expert delineations as reference and compare the applicability of the new automatic algorithm to previously suggested methods for infarct quantification in both IR and PSIR LGE images.

## Methods

### Experimental studies

Pigs with induced myocardial infarction were included from three previous studies, one mechanistic study of myocardial infarction (n=15) [17] one cardio protection study (n=15) [18] and controls from one cardio protection study previously used for validating the original weighted algorithm for infarct quantification (n=8) [13]. All three animal studies conformed to the Guide for the Care and Use of Laboratory Animals United States National Institutes of Health (NIH Publication No.85-23, revised 1996) and were approved by the Regional Ethics Committee. The experimental protocols for each of the studies have been previously published [13, 17-18]. In short, all pigs were subjected to 40 minutes occlusion with a balloon placed after the first or the second diagonal branch of the left anterior descending artery (LAD). Myocardial infarction was imaged after four hours [13, 18], six hours [17] or seven days [17] of reperfusion with either *in-vivo* 3D IR LGE (n=20), *in-vivo* 2D PSIR LGE (n=12) and/or *ex-vivo* high resolution T1-weighted images (n=38). CMR imaging was performed on a 1.5 T Philips scanner (Philips Healthcare, Best, The Netherlands). *In-vivo* LGE images were acquired approximately 20 minutes after injection of gadolinium-based contrast agent. *Ex-vivo* high resolution (0.5x0.5x0.5 mm) T1-weighted images were acquired covering the entire left ventricle (LV) with the explanted hearts placed in plastic containers and the ventricles filled with balloons containing deuterated water. For *ex-vivo* imaging, a gadolinium-based contrast agent was administered 15 minutes prior to administration of a potassium chloride bolus. Six pigs with MI were imaged, both *in-vivo* and *ex-vivo*, after seven days of reperfusion and following *ex-vivo* imaging, hearts were sliced into five mm slices and incubated in triphenyltetrazolium-chloride (TTC) for five minutes. The slices were subsequently photographed on both apical and basal sides for infarct analysis.

### Patient population

Patients with first time ST-elevation myocardial infarction (STEMI) treated with percutaneous coronary intervention (PCI) were included from the recently published clinical cardioprotection trials CHILL-MI [19] (n=58) and MITOCARE [20] (n=66). Patients underwent CMR imaging within 2-6 days following acute MI treated with PCI. Inclusion and exclusion criteria for each of the clinical trials have been previously published [19, 21]. In short, all patients had clinical signs of acute MI defined as clinical symptoms and ECG signs consistent with ST-elevation infarction or new onset of left bundle branch block (LBBB), were  $\geq$  18 years old and had symptom duration of less than



6 hours. Patients with a history of previous myocardial infarction or history of coronary revascularization were excluded. Both studies [19-20] from which patients were recruited were approved by the institutional review boards/Ethics Committees, and all patients provided written informed consent. All CMR examinations were performed on 1.5 T scanners from Philips (Philips Healthcare, Best, The Netherlands), Siemens (Siemens Healthcare, Erlangen, Germany) or GE (GE Healthcare, Waukesha, WI, USA). For infarct assessment, LGE images covering the entire LV were acquired approximately 15 minutes after injection of the gadolinium-based contrast agent. The LGE-images were acquired using a magnitude inversion-recovery (IR) or phase sensitive inversion recovery (PSIR) gradient-recalled echo sequence with a slice thickness of 8 mm with no slice gap [4]. In-plane resolution was typically 1.5 x 1.5 mm. Inversion time was manually adjusted to null the signal of viable myocardium. Surface coil intensity correction was not mandatory across vendors and sites.

This study included patients who had undergone CMR examination with LGE magnitude IR images (n=75) or paired LGE magnitude and phase sensitive IR images (n=49). Image quality was assessed as (1) poor, (2) acceptable or (3) good, where acceptable and good images were considered for this study.

## Image analysis

All images were analyzed using the software Segment (<http://segment.heiberg.se>) [22]. Endocardial and epicardial borders were traced manually with exclusion of the papillary muscles. Infarct size was expressed as % of left ventricular mass (LVM).

In LGE images, infarct expert delineation was performed using the weighted method based on 1.8SD [13], with manual corrections where needed. Hypointense regions within the region of gadolinium enhancement were considered to be MVO [23] and were included in the analysis as 100 % infarction. In the experimental *in-vivo* data, LGE images were delineated with the same method as for the patients by one observer (RJ with 5 years of CMR experience). In the patient data, delineation of each data set was performed by one of three primary observers (HE, MC and HA with 14, 15 and 20 years of experience, respectively) in a core lab setting (Imacor AB, Lund Sweden) with a quality control of the delineations by a second observer for each case. Different opinions for the delineations were resolved in consensus between all three observers when necessary. In a subset of 17 patients a second-observer delineation was performed using the same endocardial and epicardial borders to evaluate inter-observer

variability of the expert delineation (MC vs. HE). In T1-weighted images, infarct delineation was performed using a threshold of 8SD from remote [13], with manual corrections where needed (RJ or HE). Hypointense regions were considered to be MVO and included in the infarct delineation. In TTC images the myocardial infarction was delineated manually by one observer as the non TTC-stained parts of the myocardium (RJ).

## Automatic segmentation of MI

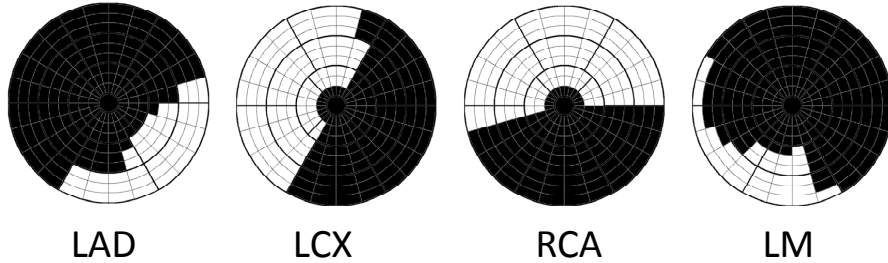
The automatic algorithm for MI quantification was implemented and incorporated in the freely available software Segment [22]. The new automatic algorithm, EWA, is based on three major principles: Expectation Maximization for intensity classification, weighted summation of infarct size to account for partial volume effects according to pixel intensity and *a priori* information utilized for pre and post processing. A maximal extent model of the perfusion territories [15] was used as *a priori* information and was defined by consensus between three experienced observers for each culprit artery (**Figure 1**). The user supplies the EWA algorithm with information on culprit artery and indicates the rotation of the left ventricle by the inferior and anterior right ventricular insertion points. The EWA algorithm consists of six steps:

- 1) Surface coil intensity correction
- 2) Classification of myocardial intensities by means of an EM-algorithm
- 3) Segmentation of infarct region by means of a level set method
- 4) Inclusion of microvascular obstruction
- 5) Post processing to exclude artifacts
- 6) Calculation of the infarct size by weighting the pixels based on their intensity.

In *step 1*, surface coil intensity correction was applied as a second order linear correction to be able to account for intensity gradient proportional to the squared coil distance and was based on the intensities in the blood pool and remote myocardium. The remote myocardium was defined by using the maximal extent model of the culprit artery [15].

In *step 2*, an EM-algorithm [14] was used to classify myocardial pixel intensities as representative of normal myocardium or infarct. The EM-algorithm iteratively refined an initial classification to find the maximum likelihood estimate of the mean and standard deviation for the Gaussian intensity distributions of normal myocardium and MI. The initial classification was defined as a small MI of 10% by a pure threshold at the 90th percentile of the intensity histogram. To increase the stability of the EM-algorithm, a constrained version was applied in which pixel intensities below the 5th percentile were kept

## Maximal extent model of perfusion territories



**Figure 1- Maximal extent model of perfusion territories**

Bulls-eye representation of maximal extent model of the perfusion territories of left anterior descending artery (LAD), left circumflex artery (LCx), right coronary artery (RCA), and left main artery (LM). Models for LAD, LCX and RCA were defined in consensus by three experienced observers in an extended 17- segment AHA model and models for LM were defined from the models of LAD, LCX and RCA. The 17-segment model is extended to three slices in each of the basal, mid-ventricular and apical zones and 24 sectors in each slice. Black sectors are included in the maximal extent model. The septal part of the left ventricle was represented in the left of the bulls-eye plot, the lateral part in the right, anterior part in the top, inferior part in the bottom, the apical slices in the center and the basal slices in the outer part of the bulls-eye plot.

classified as normal myocardium and pixel intensities above the 95th percentile were kept classified as MI. The optimal intensity threshold was then defined as the intensity for which there was an equal probability of being representative of normal myocardium and MI calculated from the Gaussian distributions with mean and standard deviation estimated by the EM-algorithm.

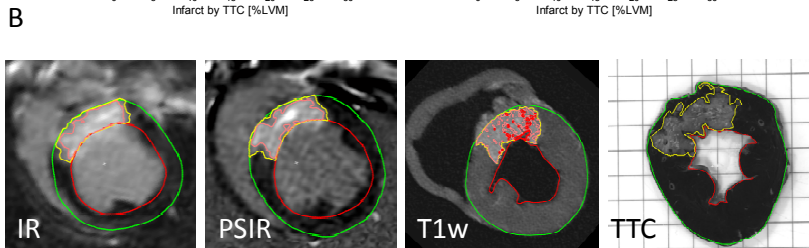
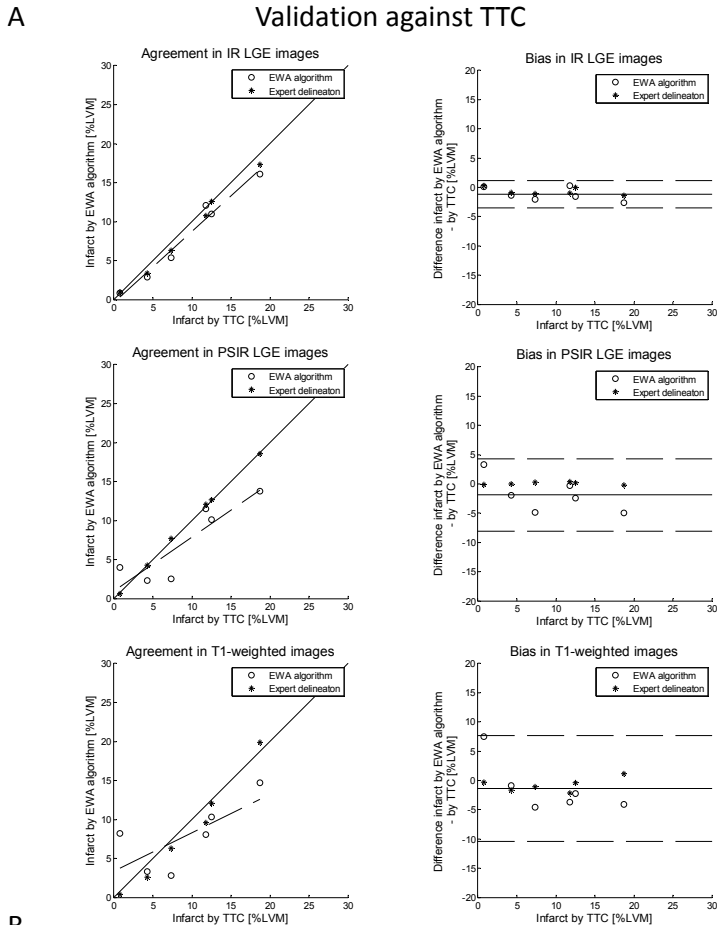
In *step 3*, the MI region was segmented using a fast level set method [24] in which the main driving force was what pixels to include or exclude based on intensity with parameters for smoothing as previously implemented by Heiberg *et al.* [13] for quantification of infarction. The main driving force in a level set method is called the speed image and should be defined to be positive for pixels to include and negative for pixels to exclude in the segmentation. Therefore, the speed image was set to a linear function with zero at the optimal threshold defined by EM, 1 at the maximal myocardial intensity and subsequently negative values for intensities below the optimal threshold.

In *step 4*, MVO was detected by means of a flood fill algorithm and morphological operations. Microvascular obstruction is characterized by regions of low intensity within the MI and might not have been detected as MI by thresholding. In the EWA algorithm MVO was detected slice by slice as holes in the infarct region by using a flood fill algorithm as suggested by Heiberg *et al.* [13] in combination with morphological closing as

suggested by Hsu *et al.* [11]. First a flood fill algorithm was used to detect dark pixels as MVO if totally surrounded by pixels segmented as infarct or connected to the endocardial border. Next a morphological closing operation was performed by first applying a dilation operation with a 3-by-3 pixel cross shaped kernel to close small gaps in the infarct segmentation. Then, the flood fill algorithm was reapplied to find any holes arising from the morphological closing before performing the erosion operation.

In *step 5*, post processing of the MI segmentation was performed in two steps: removing pixels classified as MI outside the culprit region and removing small isolated regions classified as MI. By using the same maximal extent model as for surface coil correction, bright regions outside the culprit artery region could be removed from the MI segmentation. Regions segmented as MI which were smaller than 1.5 cm<sup>3</sup> were removed regardless of location if not comprising more than 1% of the left ventricular mass or if being the only region of MI.

In *step 6*, the final step, the MI size was calculated by a weighted summation, where each pixel within the MI was weighed according to its intensity to account for partial volume effects. The weight represented the amount of infarcted cells within the pixel and hence in normal myocardium the weight should be 0 and in pixels with the maximal intensity the weight should be set to 1. The weight



**Figure 2 Validation against TTC:**

A) Scatter plots (left column) and Bland-Altman plots (right column) of infarct size expressed as % of left ventricular mass (%LVM) for the EWA algorithm against infarct size by triphenyltetrazolium-chloride (TTC) in pigs with myocardial infarction imaged after seven days ( $n=6$ ) with in-vivo magnitude inversion recovery LGE images (IR, top row), in-vivo phase sensitive inversion recovery LGE images (PSIR, middle row) and ex-vivo high resolution T1-weighted images (T1w, bottom row). Left column: solid line = line of identity; dashed line = regression line. Right column: solid line = mean bias; dashed line = mean  $\pm$  two standard deviations.

B) Infarct segmentation by the EWA algorithm in one pig shown in one slice of in-vivo IR LGE, in-vivo PSIR LGE, ex-vivo high resolution T1w and corresponding TTC-stained slice. Infarct segmentation by the EWA algorithm and by manual delineation in TTC images is shown in yellow. For the automatic EWA segmentation the core of the infarct is shown in pink and microvascular obstruction is shown as the red line within the infarct. Endocardium is delineated in red and epicardium in green.

	Magnitude IR LGE images				Phase sensitive IR LGE images				PSIR vs IR	
	Bias to reference delineation [%LVM]	Regression R-value	DSC full extent	DSC core extent	Bias to reference delineation [%LVM]	Regression R-value	DSC full extent	DSC core extent	Bias [%LVM]	Regression R-value
EWA algorithm	-3 ± 5	0.89	0.82 ± 0.14	0.81 ± 0.15	-1 ± 5	0.88	0.82 ± 0.17	0.79 ± 0.15	1 ± 4	0.92
Original weighted algorithm	-7 ± 8	0.68	0.70 ± 0.32	0.67 ± 0.32	*	*	*	*	*	*
EM threshold	6 ± 7	0.88	-	0.67 ± 0.14	6 ± 8	0.86	-	0.68 ± 0.14	0 ± 6	0.91
2 SD threshold	7 ± 7	0.85	-	0.69 ± 0.15	8 ± 6	0.86	-	0.70 ± 0.13	1 ± 5	0.94
3 SD threshold	0 ± 7	0.81	-	0.70 ± 0.21	-2 ± 7	0.79	-	0.70 ± 0.19	-2 ± 4	0.94
5 SD threshold	-8 ± 8	0.68	-	0.50 ± 0.33	-13 ± 10	0.38	-	0.36 ± 0.31	-4 ± 6	0.81
FWHM(min) threshold	-8 ± 9	0.54	-	0.58 ± 0.20	9 ± 12	0.47	-	0.69 ± 0.17	18 ± 12	0.44
FWHM(remote) threshold	**	**	-	**	-8 ± 7	0.74	-	0.66 ± 0.19	**	**
Otsu threshold	-8 ± 11	0.50	-	0.50 ± 0.32	10 ± 15	0.46	-	0.64 ± 0.20	18 ± 17	0.35

**Table 1- Bias and regional agreement in paired IR and PSIR LGE images from multi-center patient studies:** Bias as % of left ventricular mass (%LVM), regression R-value and regional agreement by DSC to expert delineation for the EWA algorithm, the original weighted algorithm [13] and the threshold method of EM, 2SD, 3SD and 5SD from remote, and FWHM from minimum intensity [8], FWHM from mean intensity in remote [11] and Otsu's threshold [25] in paired magnitude inversion recovery (IR) and phase sensitive inversion recovery (PSIR) images (n=49) and bias and regression R-value for PSIR vs IR LGE images. \* the original weighted algorithm by Heiberg et al. [13] was developed for IR images and therefore only applied in IR images. \*\* the FWHM remote threshold was developed for PSIR images as part of the FACT algorithm by Hsu et al. [11] and therefore only applied in PSIR images.

for each pixel was calculated as a linear function from weight 0 at the mean intensity of the remote myocardium to weight 1 at the 90th percentile of the intensities within the MI.

The maximal extent model of the culprit artery was needed for the intensity correction in step 1 and the first part of the post processing in step 5. However, the maximal extent model could not be applied in experimental studies where the anatomy differs and therefore the algorithm was used without the use of maximal extent model and user input of insertion points and culprit artery in the experimental part of this study.

## Statistical analysis

Experimental validation: Infarct size by the EWA algorithm and infarct size by the expert delineation in *in-vivo* IR, *in-vivo* PSIR and *ex-vivo* high resolution T1-weighted images was compared to infarct size by TTC for myocardial infarction imaged seven days after reperfusion. Infarct size by the EWA algorithm in *in-vivo* IR, *in-vivo* PSIR and *ex-vivo* high resolution T1-weighted images was compared to infarct size by expert delineation in *ex-vivo* high resolution T1-weighted images regardless of timing of imaging. Comparisons were performed using Bland-Altman bias (mean ± standard deviation) and linear regression analysis (correlation coefficient).

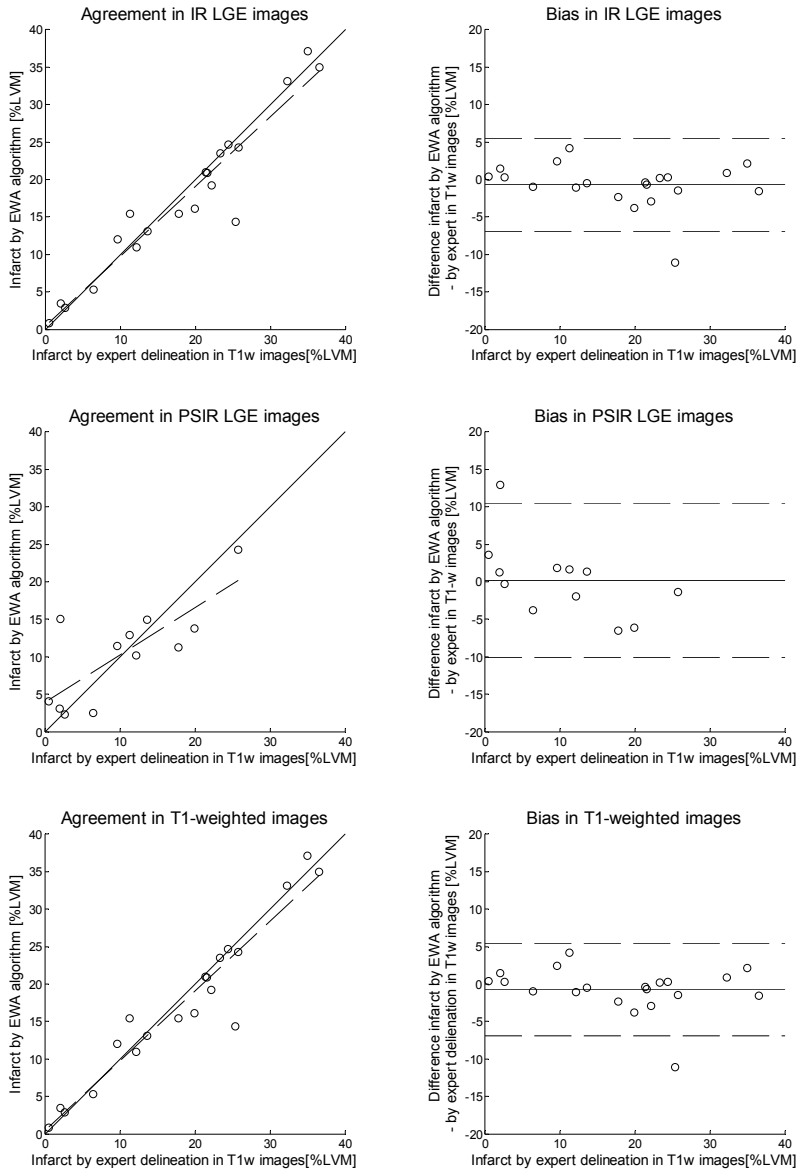
Applicability in patient data: Infarct size by the EWA algorithm was compared to infarct size by expert delineation using Bland-Altman bias (mean ± standard deviation) and linear regression analysis (correlation coefficient). Performance of the EWA

algorithm was compared to the original weighted algorithm by Heiberg et al. [13], and the thresholds of EM, 2, 3 and 5SD from remote, FWHM from minimum intensity as implemented by Amado et al. [8], FWHM from remote intensity as implemented by Hsu et al. [11] and Otsu's threshold [25]. Comparison was performed in paired IR and PSIR LGE images using bias and linear regression analysis with expert delineation as reference. Regional agreement with expert delineation was evaluated using Dice Similarity Coefficient (DSC) [26] for both the full extent of the infarct and the core of the MI as represented if no weighting had been used.

## Results

Infarct size by TTC was  $9 \pm 6$  %LVM (n=6) and infarct size by the EWA algorithm was  $8 \pm 6$  %LVM in *in-vivo* IR LGE images,  $8 \pm 5$  %LVM in *in-vivo* PSIR LGE images and  $8 \pm 4$  %LVM in *ex-vivo* high resolution T1-weighted images (T1w). The bias compared to TTC was  $-1 \pm 1$  %LVM (R=0.99) in IR images,  $-2 \pm 3$  %LVM (R=0.88) in PSIR images and  $-1 \pm 4$  %LVM (R=0.75, p=0.08) in T1w images (**Figure 2**). The infarct size by expert delineation in the same six animals was  $9 \pm 6$  %LVM in IR and PSIR images and  $9 \pm 7$  %LVM in T1w images, with a bias to TTC of  $-1 \pm 1$  %LVM (R=0.999) in IR images,  $0 \pm 0$  %LVM (R=0.994) in PSIR images and  $-1 \pm 1$  %LVM (R=0.99) in T1w images (**Figure 2**). Bias to expert delineation in T1w images was for the EWA algorithm  $-1 \pm 3$  %LVM (R=0.96, n=20) in IR LGE images,  $0 \pm 5$  %LVM (R=0.77, n=12) in PSIR LGE images and -

## Validation against ex-vivo high resolution T1-weighted images



**Figure 3- Validation against ex-vivo high resolution T1-weighted images:**

Scatter plots (left column) and Bland-Altman plots (right column) of infarct size expressed as % of left ventricular mass (%LVM) for the EWA algorithm against infarct size by expert delineation in ex-vivo high resolution T1-weighted images (T1w). Validation in in-vivo magnitude inversion recovery (IR, top row,  $n=20$  pigs), in-vivo phase sensitive inversion recovery (PSIR, middle row,  $n=12$ ) and ex-vivo high resolution T1-weighted images (T1w, bottom row,  $n=38$ ). Left column: solid line = line of identity; dashed line = regression line. Right column: solid line = mean bias; dashed line = mean  $\pm$  two standard deviations.

$1 \pm 4$  %LVM ( $R=0.96$ ,  $n=38$ ) in T1w images (Figure 3).

Infarct size by the EWA algorithm was  $15 \pm 8$  %LVM in IR images ( $n=124$ ) with a bias of  $-2 \pm 6$  %LVM ( $R=0.81$ ) compared to the expert delineation. In patients with paired IR and PSIR images ( $n=49$ ) infarct size by the EWA algorithm was  $17 \pm 10$  %LVM in both IR and PSIR images with a bias of  $-1 \pm 5$  %LVM ( $R=0.89$ ) in both IR and PSIR images (Figure 4). The bias and correlation between expert delineation of infarct size and the EWA algorithm, the original weighted algorithm, thresholds by EM, 2SD, 3SD and 5SD from remote, FWHM from minimum intensity [8], FWHM from mean intensity in remote [11] and Otsu's threshold [25] are summarized in Table 1 and Figure 5. Inter-observer variability of infarct size by expert delineation was  $0 \pm 1$  %LVM ( $R=0.99$ ).

## Discussion

This study has presented a new automatic algorithm, the EWA algorithm, for MI quantification based on intensity classification by Expectation Maximization (EM) and weighting each pixel according to its intensity to account for partial volume effects. The EWA algorithm was validated experimentally and applied in multi-center, multi-vendor patient data with a low bias and high regional agreement in both IR and PSIR LGE images. The performance of the EWA algorithm was found superior to several previously described methods for MI quantification and the EWA algorithm was successfully applied to high resolution T1-weighted images, showing the ability of the EWA algorithm to adapt to different image quality.

## Experimental validation

The EWA algorithm was validated against TTC with bias similar to FWHM from minimum intensity as suggested by Amado *et al.* [8] ( $4.1 \pm 1.1$  %LVM,  $R=0.94$ ) and the FACT algorithm by Hsu *et al.* [11] (1.9% LVM,  $R=0.96$ ). The EWA algorithm was also validated against *ex-vivo* high resolution T1-weighted images in a larger cohort with bias comparable to the original weighted algorithm by Heiberg *et al.* [13] ( $-0.3 \pm 1.3$  %LVM). The EWA algorithm was validated against TTC in pigs with myocardial infarction imaged after seven days of reperfusion and against *ex-vivo* high resolution T1-weighted images in pigs with myocardial infarction imaged after four hours, six hours or seven days of reperfusion.

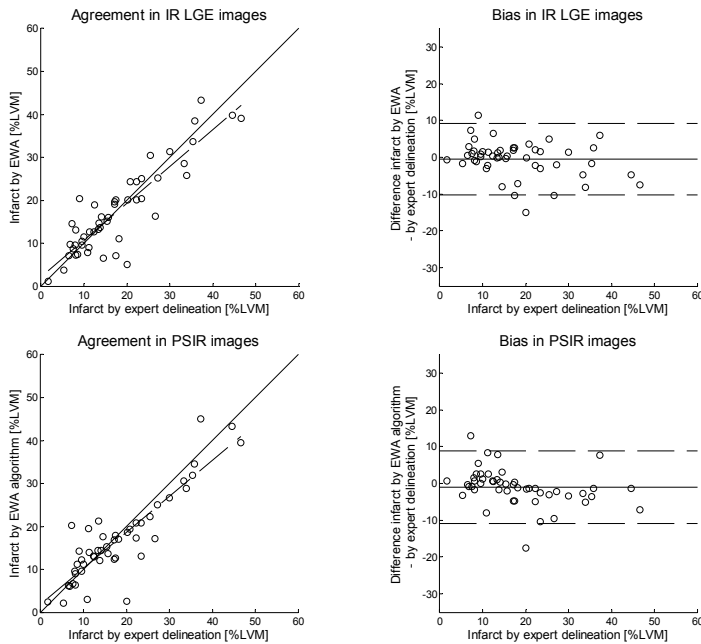
Hyperenhancement in CMR has been shown to overestimate acute MI in comparison to TTC [17, 27-28] and Jablonowski *et al.* [17] showed an overestimation by CMR after 6 hours of reperfusion which was not seen after seven days of reperfusion. The overestimation in the acute setting was explained by an increased extracellular volume adjacent to the infarct which was not seen after seven days of reperfusion [17]. Thus, TTC can be used as reference in myocardial infarction imaged seven days after reperfusion but another reference was needed for quantification of acute MI. *Ex-vivo* high resolution T1-weighted images and inversion recovery LGE are based on the same principle of hyperenhancement proportional to the distribution of the gadolinium based contrast agent in the extracellular volume. *Ex-vivo* imaging enables highresolution imaging and therefore *ex vivo* high resolution T1-weighted imaging can be used as reference for *in-vivo* IR and PSIR LGE in both the acute and chronic setting.

However, neither TTC nor *ex-vivo* high resolution T1-weighted images can be used for validation in patient studies. In this study, expert delineation was chosen as the reference for MI quantification in patients, performed by using the original weighted algorithm by Heiberg *et al.* [13] followed by manual corrections and consensus reading. The expert delineation was validated against TTC demonstrating a lower bias compared to manual delineation in the study by Amado *et al.* [8] and Hsu *et al.* [11] ( $8.6 \pm 1.9$ % LVM,  $R=0.69$  and 5.4%,  $R=0.96$ , respectively). Interobserver variability was analyzed in patients in a core lab setting and showed a lower bias and variability compared with previous studies by Flett *et al.* [9] and McAlindon *et al.* [29]. Thus, the expert delineation was used as reference in the patient population.

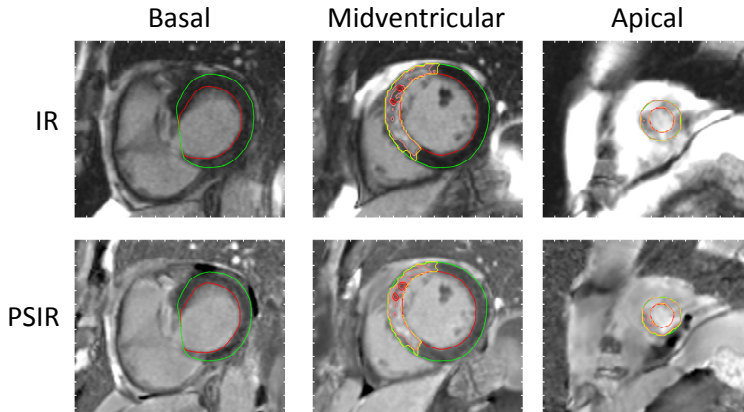
## Applicability in multi-center patient data

The EWA algorithm was applied in 124 patients from multi-center, multi-vendor studies with bias to expert delineation similar to the FACT algorithm by Hsu *et al.* [12] which was evaluated in 20 patients from a single center (3.8 %LVM,  $R=0.95$ ). Heiberg *et al.* [13] found a lower bias for the original weighted algorithm ( $0.3 \pm 2.7$  %LVM) in a two-vendor, single-center study of 40 patients. However, in the present study, the performance of the EWA algorithm was compared to the original weighted algorithm [13] and a higher bias and variability was found for the original weighted algorithm in the current multi-center, multi-vendor study than in the original study [13]. Similarly,

## A Applicability in paired IR and PSIR LGE images from patients



## B



**Figure 4- Applicability in paired IR and PSIR LGE images from patients in multi-center, multi-vendor studies:**

A) Scatter plots (left column) and Bland-Altman plots (right column) of infarct size expressed as % of LVM for the EWA algorithm against infarct size by expert delineation in 49 patients from multi-center studies with paired magnitude inversion recovery (IR, top row) and phase sensitive inversion recovery LGE images (PSIR, bottom row). Left column: solid line = line of identity; dashed line = regression line. Right column: solid line = mean bias; dashed line = mean  $\pm$  two standard deviations.

B) Typical segmentation by the EWA algorithm in one patient with paired IR (top row) and PSIR images (bottom row). The automatic EWA segmentation of the infarct is shown in yellow, the core of the infarct is shown in pink and microvascular obstruction is shown as the red line within the infarct. Endocardium is delineated in red and epicardium in green.

variability was increased for the threshold by FWHM from minimum intensity and n-SD from remote in comparison to the validation against TTC by Amado *et al.* [8] and in contrast to the study by Hsu *et al.* [11] an underestimation was seen for the threshold of FWHM from remote. The changes in bias and variability seen in the current multi-center, multi-vendor patient study compared to previous validations in experimental studies [8, 11, 13] and single-center patient studies [12-13] underlines the importance of using multi-center, multi-vendor patient data. Multi-center, multi-vendor patient data has a larger variability in image quality and thus the automatic algorithm is faced with a larger challenge which may not have been accounted for in the algorithm if designed and validated for single-center patient data or experimental data.

Additionally, infarct validation needs to be performed in both magnitude IR and PSIR images since both are used in clinical routine. Stirrat *et al.* [10] recently showed a significant bias of infarct size in paired magnitude IR and PSIR images for n-SD from remote and FWHM from minimum intensity. Based on their findings we compared infarct size in paired IR and PSIR images to expert delineation in 49 patients for the EWA algorithm, threshold methods of EM, 2, 3, and 5 SD, FWHM from minimum intensity and Otsu's threshold. There was a large bias between IR and PSIR images for the threshold of FWHM from minimum intensity and Otsu's threshold with underestimation in IR and overestimation in PSIR images. Bias between IR and PSIR for 2, 3 and 5 SD was lower in this study than in the study by Stirrat *et al.* [10] (-3 %LVM, -4 %LVM and -5 %LVM, respectively) and is possibly explained by different definitions of remote region. In the present study the remote region was defined from the *a priori* maximal extent model for each culprit artery. In the study by Stirrat *et al.* [10] care was taken to manually define a large remote region, however, infarct size in controls without myocardial infarction was found as 14 %LVM by 2SD and 9 %LVM by 3SD instead of the theoretically defined 2 % and 0.1 %. The difficulty in defining a remote region representative

## Conclusion

We have developed a new automatic algorithm, the EWA algorithm, for quantification of myocardial

volume effects. The EWA algorithm performed well for both magnitude IR and PSIR LGE images when validated in experimental studies against TTC and *ex-vivo* high resolution T1-weighted images, and when applied in multi-center, multi-vendor

of normal myocardium is also shown by high variability of 2SD in inter- and intra observer variability and test-retest repeatability found by both Flett *et al.* [9] and McAlindon *et al.* [29]. By using the EWA algorithm there is no need for manual definition of remote regions and the EWA algorithm showed a lower variability and higher regional agreement than any other of the methods and a low bias and variability between IR and PSIR images.

## Limitations

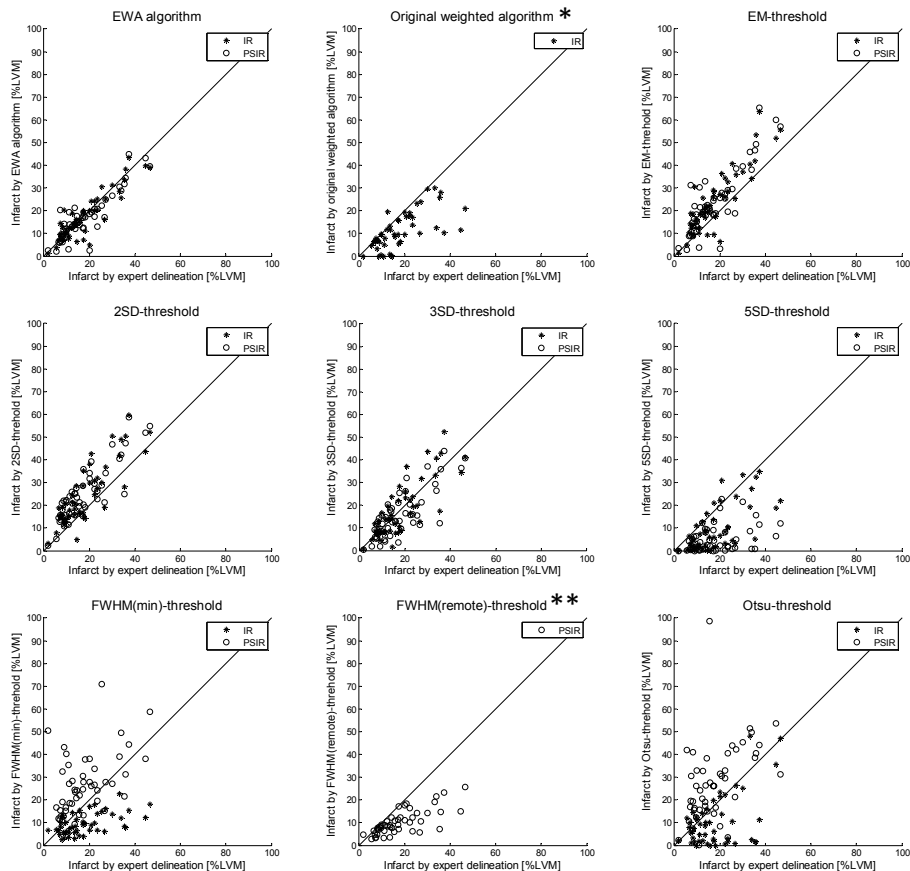
The EWA algorithm was applied in multi-center, multi-vendor patient data from clinical trials of first time STEMI and experimental studies of a single infarction and the EWA algorithm was developed for single vessel myocardial infarction. For multi-vessel myocardial infarction or multiple infarctions over time the algorithm can however be used without the *a priori* information of culprit artery models. The algorithm would then not be able to apply the intensity correction and would need further validation for multi-vessel disease. For other types of myocardial fibrosis such as in the situation of hypertrophic cardiomyopathy and myocarditis both *a priori* information and post processing might need to be adjusted and would require additional validation for these groups of patients. However, the EWA algorithm was applied in the experimental data without the use of *a priori* information on culprit artery due to differences in anatomy and showed a low bias in IR and PSIR LGE images and *ex-vivo* high resolution T1-weighted images. The low bias found in T1-weighted images as well as in IR and PSIR LGE images shows the ability of the EWA algorithm to assess infarct size in a wide range of settings with a variety of different imaging strategies. The need for manual corrections was not assessed, however, considering the lower bias and higher regional agreement than for the original weighted algorithm less manual corrections would probably be needed. Especially for quantification in *ex-vivo* high resolution T1-weighted images, time will be saved by the limited amount of user input in comparison to definition of remote regions in all 0.5 mm slices covering the left ventricle.

infarction in LGE images based on combining an intensity classification by Expectation Maximization (EM) with a pixel intensity weighting approach to account for partial

patient data. Thus, using EM and a weighted approach as with the EWA algorithm, may serve as a candidate for a clinical standard in quantifying myocardial infarction.



## Applicability in paired IR and PSIR LGE images compared to previously suggested methods



**Figure 5- Applicability in paired IR and PSIR LGE images from multi-center patient studies compared to previously suggested methods for MI quantification:**

Scatter plots of infarct size expressed as % of left ventricular mass (% LVM) against infarct size by expert delineation in 49 patients for the EWA algorithm, the original weighted algorithm [13] and the threshold method of Expectation Maximization (EM) [14] (top row), 2SD, 3SD and 5SD from remote (middle row), and FWHM from minimum intensity [8], FWHM from mean intensity in remote [11] and Otsu's threshold [25] (bottom row) in paired magnitude inversion recovery (IR) and phase sensitive inversion recovery (PSIR) LGE images. Solid lines = line of identity. \* the original weighted algorithm by Heiberg et al. [13] was developed for IR images and therefore only applied in IR images. \*\* the FWHM remote threshold was developed for PSIR images as part of the FACT algorithm by Hsu et al. [11] and therefore only applied in PSIR images.

### Abbreviations

CMR: cardiovascular magnetic resonance, DSC : Dice similarity coefficient, EM : expectation maximization, FWHM : full width half maximum, IR : inversion recovery, KCl : potassium chloride, LAD : left anterior descending artery, LBBB: left bundle branch block, LCx : left circumflex artery,

LGE : late gadolinium enhancement, LM : left main artery, MI : myocardial infarction, PCI: percutaneous coronary intervention, PSIR: phase sensitive inversion recovery, RCA : right coronary artery, SD: standard deviations, STEMI: ST-elevation myocardial infarction, %LVM : percent of left ventricular mass

## Acknowledgements

This study has been funded by the Swedish Research Council (2011-3916, 2012-4944), The Swedish Heart and Lung Foundation, The Medical Faculty of Lund University, Sweden, and Region of Scania, Sweden. Funding for the CHILL-MI trial was received from Philips Healthcare. Funding for the MITOCARE trial was received from the European Commission within the 7<sup>th</sup> Framework Programme for RTD – Project MITOCARE – Grant Agreement HEALTH-2010-261034.

## Author details

<sup>1</sup> Clinical Physiology, Dept. of Clinical Sciences Lund, Lund University, Skåne University Hospital, Sweden, <sup>2</sup>Dept. of Biomedical Engineering, Faculty of Engineering, Lund University, Lund, Sweden, <sup>3</sup>Laboratory of Medical Informatics, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece, <sup>4</sup>Dept. of Cardiology B, Oslo, University Hospital Ullevål and Faculty of Medicine, University of Oslo, Oslo, Norway, <sup>5</sup>UMR 7339 CRMBM, Aix-Marseille University, Marseille, France, <sup>6</sup>Department of Cardiology, Hospital Nord, Marseille, France, <sup>7</sup>Department of Cardiology, Medical University of Innsbruck, Innsbruck, Austria, <sup>8</sup>Dept. of Cardiology, Lund University, Lund, Sweden

Corresponding author: Einar Heiberg, Department of Clinical Physiology, Skåne University Hospital, Lund, SE 221 85 Lund, Sweden

## Authors' contributions

JT developed and implemented the EWA algorithm, analyzed and interpreted segmentation results, drafted the manuscript and conceived the study. HE designed the experimental part of the study, contributed to the overall design of the study and the automatic algorithm and performed manual delineations. RJ and MC contributed to the design of the study and the automatic algorithm and performed manual delineations. AA contributed to the design of the study and the automatic algorithm and provided in-depth CMR knowledge needed for algorithm development. PH, AJ, FK, BM, DE and DA were responsible for data collection in the multi-center studies. HA contributed to the design of the automatic algorithm, performed manual delineations and conceived the study. EH contributed to the design of the automatic algorithm and conceived the study. All authors revised the manuscript, and have read and approved the final version of the manuscript.

## Competing interests

EH is the founder of Medviso AB, Lund, Sweden, which sells a commercial version of Segment. JT is employed by Medviso AB on a part-time basis. HA is a share-holder of Imacor AB, Lund, Sweden,

which performs core lab analysis of CMR images. HA, HE and MC have been employed by Imacor AB on a part-time basis. The other authors declare that they have no competing interests.

## References

1. Kim, R.J., et al., Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. *Circulation*, 1999. **100**(19): p. 1992-2002.
2. Fieno, D.S., et al., Contrast-enhanced magnetic resonance imaging of myocardium at risk: distinction between reversible and irreversible injury throughout infarct healing. *J Am Coll Cardiol*, 2000. **36**(6): p. 1985-91.
3. Wesbey, G.E., et al., Effect of gadolinium-DTPA on the magnetic relaxation times of normal and infarcted myocardium. *Radiology*, 1984. **153**(1): p. 165-9.
4. Simonetti, O.P., et al., An improved MR imaging technique for the visualization of myocardial infarction. *Radiology*, 2001. **218**(1): p. 215-23.
5. Kellman, P., et al., Phase-sensitive inversion recovery for detecting myocardial infarction using gadolinium-delayed hyperenhancement. *Magn Reson Med*, 2002. **47**(2): p. 372-83.
6. Schulz-Menger, J., et al., Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) board of trustees task force on standardized post processing. *J Cardiovasc Magn Reson*, 2013. **15**: p. 35.
7. Bondarenko, O., et al., Standardizing the definition of hyperenhancement in the quantitative assessment of infarct size and myocardial viability using delayed contrast-enhanced CMR. *J Cardiovasc Magn Reson*, 2005. **7**(2): p. 481-5.
8. Amado, L.C., et al., Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. *J Am Coll Cardiol*, 2004. **44**(12): p. 2383-9.
9. Flett, A.S., et al., Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance. *JACC Cardiovasc Imaging*, 2011. **4**(2): p. 150-6.
10. Stirrat, J., et al., Influence of phase correction of late gadolinium enhancement images on scar signal quantification in patients with ischemic and non-ischemic cardiomyopathy. *J Cardiovasc Magn Reson*, 2015. **17**(1): p. 66.
11. Hsu, L.Y., et al., Quantitative myocardial infarction on delayed enhancement MRI. Part I: Animal validation of an automated feature analysis and combined thresholding infarct sizing algorithm. *J Magn Reson Imaging*, 2006. **23**(3): p. 298-308.
12. Hsu, L.Y., et al., Quantitative myocardial infarction on delayed enhancement MRI. Part II: Clinical application of an automated feature analysis and combined thresholding infarct sizing algorithm. *J Magn Reson Imaging*, 2006. **23**(3): p. 309-14.
13. Heiberg, E., et al., Automated quantification of myocardial infarction from MR images by accounting for partial volume effects: animal, phantom, and human study. *Radiology*, 2008. **246**(2): p. 581-8.
14. Dempster, A.P., N.M. Laird, and D.B. Rubin, Maximum Likelihood from Incomplete Data Via Em Algorithm. *Journal of the Royal Statistical Society Series B-Methodological*, 1977. **39**(1): p. 1-38.
15. Sjogren, J., et al., Semi-automatic segmentation of myocardium at risk in T2-weighted cardiovascular magnetic resonance. *J Cardiovasc Magn Reson*, 2012. **14**: p. 10.
16. Wei, D., et al., A comprehensive 3-D framework for automatic quantification of late gadolinium enhanced cardiac magnetic resonance images. *IEEE Trans Biomed Eng*, 2013. **60**(6): p. 1499-508.
17. Jablonowski, R., et al., Contrast-enhanced CMR overestimates myocardial infarction size on day 1 but not day 7 relative to TTC in a swine model: Mechanistic insights

using extracellular volume measurements. *Journal of American College of Cardiology Cardiovascular Imaging*, 2015. **In Press**.

18. Hansson, M.J., et al., Differences in the profile of protection afforded by TRO40303 and mild hypothermia in models of cardiac ischemia/reperfusion injury. *Eur J Pharmacol*, 2015. **760**: p. 7-19.
19. Erlinge, D., et al., Rapid endovascular catheter core cooling combined with cold saline as an adjunct to percutaneous coronary intervention for the treatment of acute myocardial infarction. The CHILL-MI trial: a randomized controlled study of the use of central venous catheter core cooling combined with cold saline as an adjunct to percutaneous coronary intervention for the treatment of acute myocardial infarction. *J Am Coll Cardiol*, 2014. **63**(18): p. 1857-65.
20. Atar, D., et al., Effect of intravenous TRO40303 as an adjunct to primary percutaneous coronary intervention for acute ST-elevation myocardial infarction: MITOCARE study results. *Eur Heart J*, 2015. **36**(2): p. 112-9.
21. Rationale and design of the 'MITOCARE' Study: a phase II, multicenter, randomized, double-blind, placebo-controlled study to assess the safety and efficacy of TRO40303 for the reduction of reperfusion injury in patients undergoing percutaneous coronary intervention for acute myocardial infarction. *Cardiology*, 2012. **123**(4): p. 201-7.
22. Heiberg, E., et al., Design and validation of Segment--freely available software for cardiovascular image analysis. *BMC Med Imaging*, 2010. **10**: p. 1.
23. Beek, A.M., R. Nijveldt, and A.C. van Rossum, Intramyocardial hemorrhage and microvascular obstruction after primary percutaneous coronary intervention. *Int J Cardiovasc Imaging*, 2010. **26**(1): p. 49-55.
24. Nilsson, B. and A. Heyden, A fast algorithm for level set-like active contours. *Pattern Recognition Letters*, 2003. **24**(9-10): p. 1331-1337.
25. Otsu, N., Threshold Selection Method from Gray-Level Histograms. *Ieee Transactions on Systems Man and Cybernetics*, 1979. **9**(1): p. 62-66.
26. Dice, L.R., Measures of the Amount of Ecologic Association between Species. *Ecology*, 1945. **26**(3): p. 297-302.
27. Saeed, M., et al., Reperfused myocardial infarction as seen with use of necrosis-specific versus standard extracellular MR contrast media in rats. *Radiology*, 1999. **213**(1): p. 247-57.
28. Saeed, M., et al., Magnetic resonance characterization of the peri-infarction zone of reperfused myocardial infarction with necrosis-specific and extracellular nonspecific contrast media. *Circulation*, 2001. **103**(6): p. 871-6.
29. McAlindon, E., et al., Quantification of infarct size and myocardium at risk: evaluation of different techniques and its implications. *Eur Heart J Cardiovasc Imaging*, 2015.



