Medical Imaging of the Heart: Quantitative Analysis of Three-Dimensional Echocardiographic Images

Kun Wang

Supervisors: Dr Andrew Sims, Prof. Alan Murray

Department of Medical Physics
Institute of Cellular Medicine
Newcastle University

Degree: Doctor of Philosophy

June 2011

Abstract

Accurate, reproducible determination of cardiac chamber volume, especially left ventricular (LV) volume, is important for clinical assessment, risk stratification, selection of therapy, and serial monitoring of patients with cardiovascular disease. Echocardiography is the most widely used imaging modality in the clinical diagnosis of left ventricular function abnormalities. In the last 15 years, developments in real time three-dimensional echocardiography (RT3DE) have achieved superior accuracy and reproducibility compared with conventional two-dimensional echocardiography (2DE) for measurement of left ventricular volume and function. However, RT3DE suffers from the limitations inherent to the ultrasonic imaging modality and the cost of increased effort of data handling and image analysis.

There were two aims of this research project. Firstly, it aimed to develop different new semi-automated algorithms for LV endocardial surface delineation, LV volume and EF quantification from clinical RT3DE images. Secondly, through assessing and comparing the performance of these algorithms in the aspects of accuracy and reproducibility, this project aimed to investigate what factors in real time 3D echo images influenced the performance of each algorithm, so that advantages and drawbacks of 3D echo images can be better understood.

The basic structure of the content of this thesis is as follows: Chapter 1 introduces the background and the aims of this project. Chapter 2 describes the development of the new semi-automated algorithms. Chapter 3 to Chapter 6 presents the four studies designed to assess and compare the accuracy and reproducibility of each algorithm. These studies were the balloon phantom study, the tissue-mimicking phantom study, the clinical cardiac magnetic resonance images study and the clinical contrast enhanced 3D stress echo images study. Chapter 7 summarises all these studies, draws conclusions, and describes future work.

In conclusion, it has been shown that the semi-automated algorithms can measure LV volume and EF quantitatively in clinical 3D echo images. To achieve better accuracy and reproducibility, 3D echo images should be analysed from all three dimensions.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Professor Alan Murray and Doctor Andrew Sims, for their support and supervision of the last three years of my PhD study. Without their advice and help, it would have been impossible for me to finish this project.

I also would like to give many thanks to all the members in the Medical Physics Department, Freeman Hospital in Newcastle. As a group, we always help and encourage each other. I really enjoyed working together with all of you.

Thanks George Parkin, Tong Jie and Mike Feeney, who helped me so much for producing the tissue-mimicking phantoms. Thanks Doctor Tim Irvine, who helped me for the phantom echo scanning and clinical data collection. Thanks Professor Andrew Blamire and Doctor Kieren Hollingsworth, who provided me the cardiac magnetic resonance images.

A very big thank you goes to Doctor Dingchang Zheng and his wife, Doctor Aiqing Chen. They are not just my friends, but also my teachers and models. Thank you for all your help for both my studying and my life in Newcastle.

Finally, I would like to express to my deepest love to my parents. They respected my decision of studying the PhD in Newcastle, and they have been always supporting me unconditionally in my whole life.

Table of contents

List of figures

List of tables

Chapter 1. Introduction

Echocardiography is a valuable non-invasive tool for imaging the living heart. It is based on detection of echoes produced by a beam of ultrasound (very high frequency sound) pulses transmitted into the heart (1).

As one of the most important cardiac imaging modalities, echocardiography is used to evaluate cardiac chamber volume, wall thickness, wall motion, valvular anatomy, valve motion, the proximal great vessels and the pericardium (1, 2). It is used to determine anatomic relationships and to measure cardiac function. It is a sensitive tool for detecting pericardial and pleural fluid, identifying mass lesions within and adjacent to the heart, characterizing congenital cardiac defects and diagnosing valvular and myocardial pathology (2).

Compared with other non-invasive cardiac imaging modalities, such as magnetic resonance imaging (MRI), X-ray computed tomographic imaging (CT) or nuclear medicine, echocardiography shows many inherent advantages. It is relatively inexpensive, portable, avoids ionizing radiation, and has an excellent temporal resolution (3). In medical practice, it is very effective in diagnosing a wide variety of cardiac diseases, as well as ascertaining the severity of diseases and making a prognosis after clinical treatments. Echocardiographic examination has been widespread in cardiology, and is a crucial part of the complete cardiac diagnostic cycle. By integrating with medical history, physical examination findings, electrocardiogram, thoracic radiographs and clinical laboratory findings, echocardiography offers a great opportunity to arrive at the correct diagnosis and develop an optimal therapeutic plan for each individual patient (2, 3).

1.1. History and types of echocardiography

1.1.1. History of ultrasonic imaging

Historically, the first practical realization of ultrasonic imaging was born during World War I in the quest for detecting submarines. Soon after that, these attempts were followed by echographic techniques adapted to industrial applications for non-destructive testing of metals. Before these applications, the essential to these developments was already published in the book of “The Theory of Sound” by Lord Rayleigh in 1877, and in 1880, Pierre Curie researched the piezoelectric effect, which enabled easy generation and detection of ultrasonic waves (4). The early use of ultrasound as a diagnostic tool dates back to 1942 when two Austrian brothers used transmission of ultrasound through the brain to locate tumours. In 1949, the pulse-echo system was described (4, 5). Then, 2D grey scale images were produced during the 1950s. A real time 2D grey scale image was produced in 1965 by a scanner developed by Siemens. Since the mid-1970s, electronic scanners have been available from many companies. Image quality steadily improved during the 1980s with substantial enhancements since the mid-1990s (4, 5).

At the present time, there are three types of echocardiography in routine clinical use: M-mode, two-dimensional (B-mode or real time 2D), and Doppler echocardiography. Each of them has advantages, and all are routinely used in echocardiographic examinations. The findings from one type of echocardiographic examination complements those from other examinations.

This research project mainly focuses on non-invasive quantitative measurements of heart chamber volume, especially the left ventricular volume, which requires the analysis of anatomic images. Therefore, the two types of echocardiography involved here are real time two-dimensional (RT2DE) and the newer, but not widely used real time three-dimensional echocardiography (RT3DE).

1.1.2. M-mode echocardiography

The M-mode echocardiogram yields a one-dimensional view of the cardiac structures moving over time (4). The principle is simple: the acquired echo signals are identical for

a static transducer (see 1.1.3.) and object, but if the object moves, the signal changes. In the clinical application, the echo signal changes while various tissue interfaces along the axis of the ultrasound beam are moving during the cardiac cycle (5). This phenomenon provides an image with depth vs. time, which is called M-mode imaging.

1.1.3. Real time two-dimensional echocardiography (RT2DE)

Diagnostic ultrasound employs pulsed high frequency (>20,000 Hz) sound waves. They are generated and detected by a piezoelectric crystal, which deforms under the influence of an electric field and, vice versa, induces an electric field across the crystal during deformation. This crystal is embedded in a transducer (Figure 1-1) that serves both as a transmitter and a detector.

Figure 1-1 A typical 2D phased array transducer

A 2D image can be obtained by tilting the transducer (Figure 1-2 (a)). However, a mechanical displacement or rotation of the transducer yields several practical difficulties. For example, the mechanical displacement rate must be kept constant, and the contact between the transducer and the patient's skin must not be lost. To overcome those problems, the electronic scanning array transducer was invented. An array transducer is a collection of many small identical crystals that can be excited independently by an electric field (4). In echocardiography, phased-array transducers that consist of a 2D array of crystals are commonly used. Its biggest advantage is that it has a relatively smaller cross section than other types of transducers. For applications in which the acoustic window is small, such as in cardiology where the waves can only approach the heart through the small space between ribs, a phased array transducer is favourable. Typically 64 to 94 crystals are used to change the direction of propagation of the wave by tuning the phases of the waves sent by different crystals. This way is

electronically equivalent to tilting the transducer mechanically. So in clinical application, the operator only needs to hold the transducer steadily, then a real time 2D sector image can be obtained (Figure 1-2 (b)).

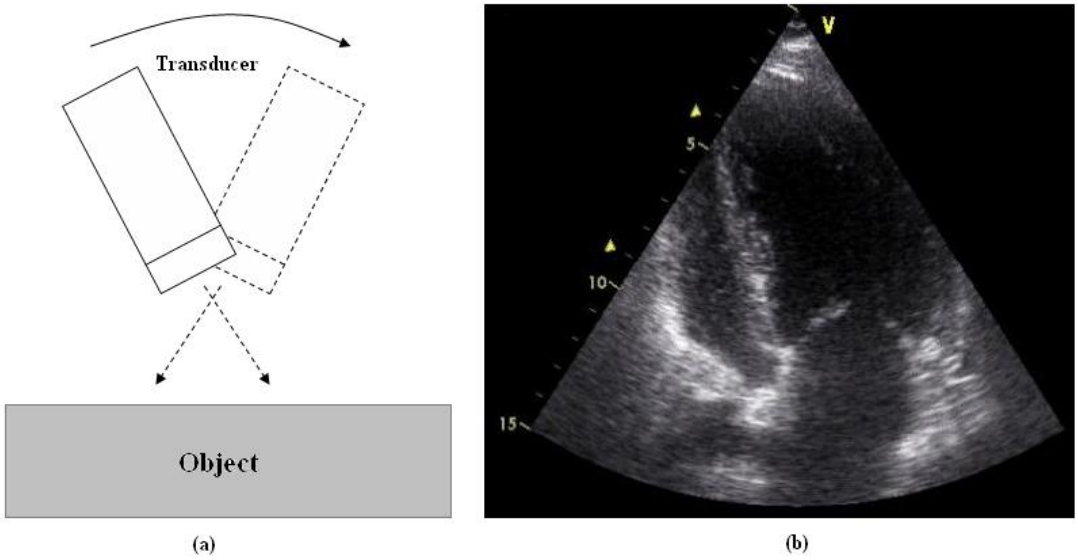


Figure 1-2 (a) 2D image acquisition can be done by tilting the transducer. (b) A RT2DE image of a normal heart in a 4-chamber view (6)

A commercial RT2D echocardiographic scanner is small and mobile. It consists of a transducer connected to a signal processing box, which displays the reconstructed images on a monitor in real time.

1.1.4. Doppler echocardiography

Doppler echocardiography is used to visualize the blood flow patterns, direction and velocity (5). It is based on detection of frequency changes (the Doppler shift) occurring as ultrasound waves reflect off blood cells moving either away from or toward the transducer. As long as the angle of incidence of the ultrasound beam and the path of blood flow is not 90 degrees, the blood flow velocity can be calculated from the frequency difference between the transmitted and received waves.

Two types of Doppler echocardiography are used clinically: pulsed wave (PW) and continuous wave (CW) (5). PW Doppler transmits short ultrasound bursts along a certain line, and reflected pulses are taken at a fixed time (Since the transmission speed is approximate constant, a fixed time equals to a fixed distance) to reconstruct the

Doppler spectrogram (4). The advantage of this method is that the blood flow velocity and direction from a specified point in a heart or blood vessel can be calculated. The main disadvantage is that the maximum velocity that can be measured is limited because the pulse repetition frequency is limited. CW Doppler uses dual crystals so that ultrasound waves can be simultaneously and continuously sent and received (4, 5). Therefore, high velocity flows can be measured. But the sampling of blood flow velocity and direction occurs all along the ultrasound beam rather than at a specified point.

Colour flow Doppler echocardiography is a form of PW Doppler which combines the M-mode and 2DE with blood flow imaging. In this technique, multiple sample volumes are analyzed along multiple scan lines. The frequency shift obtained from these many sample volumes is colour coded for direction and velocity (4, 5). In most cases, blood flow toward the transducer is coded as red and flow away is coded as blue.

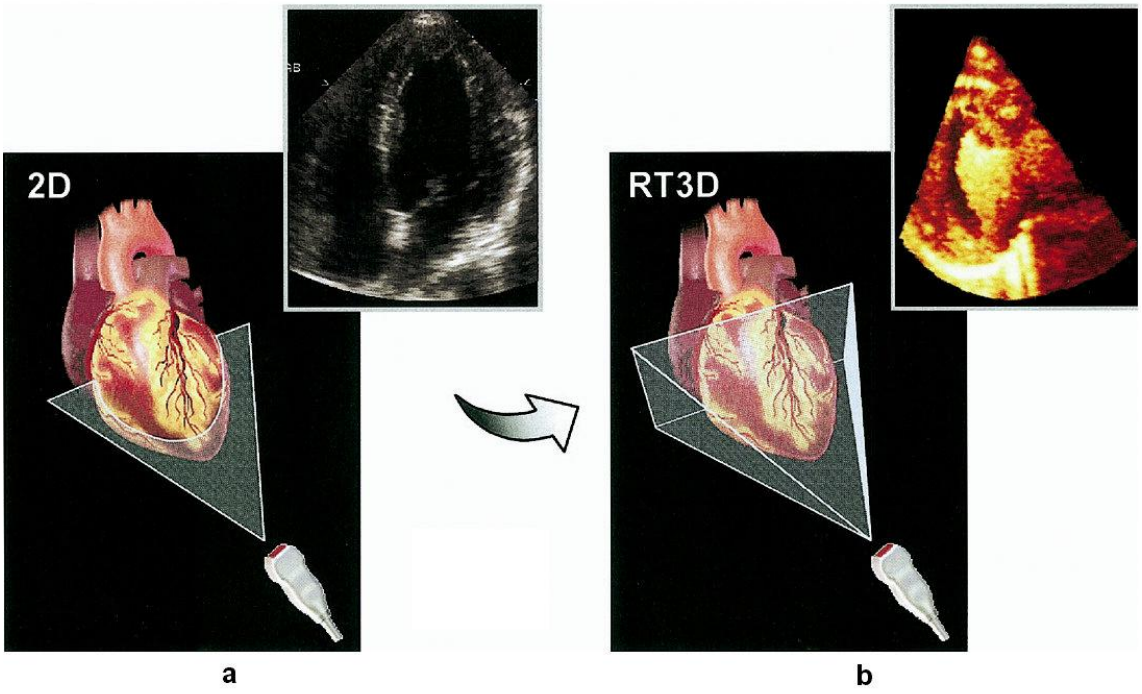


Figure 1-3 Transition from 2D to 3D imaging. (a) 2D imaging is based on scanning a single cross-sectional plane of the heart; (b) 3D imaging scans a pyramidal volume (7) covering entire left ventricle.

1.1.5. Real time three-dimensional echocardiography (RT3DE)

One of the most significant developments of the last decade in ultrasound imaging of the heart is the evolution of real time three-dimensional echocardiography (RT3DE). Different from the conventional RT2DE that scans a plan of a heart dynamically (Figure 1-3 (a)), RT3DE performs a dynamic volumetric scan that is able to cover an entire left ventricle (Figure 1-3 (b)).

Since images changed from 2D plan to 3D volume, the difficulty of image processing is also increased during imaging acquisitions. Until recently, ECG-gated 'full-volume' acquisition mode was used to capture the entire left ventricle (LV) section-by-section over several cardiac cycles (Figure 1-4 top). The major drawback of this approach was misregistration of the subvolumes manifesting itself as 'stitch artefacts' as a result of irregular heart rhythm. Respiration or any movement of the patient or transducer during image acquisition frequently leads to necessary repeated acquisitions to obtain one high-quality dataset (8).

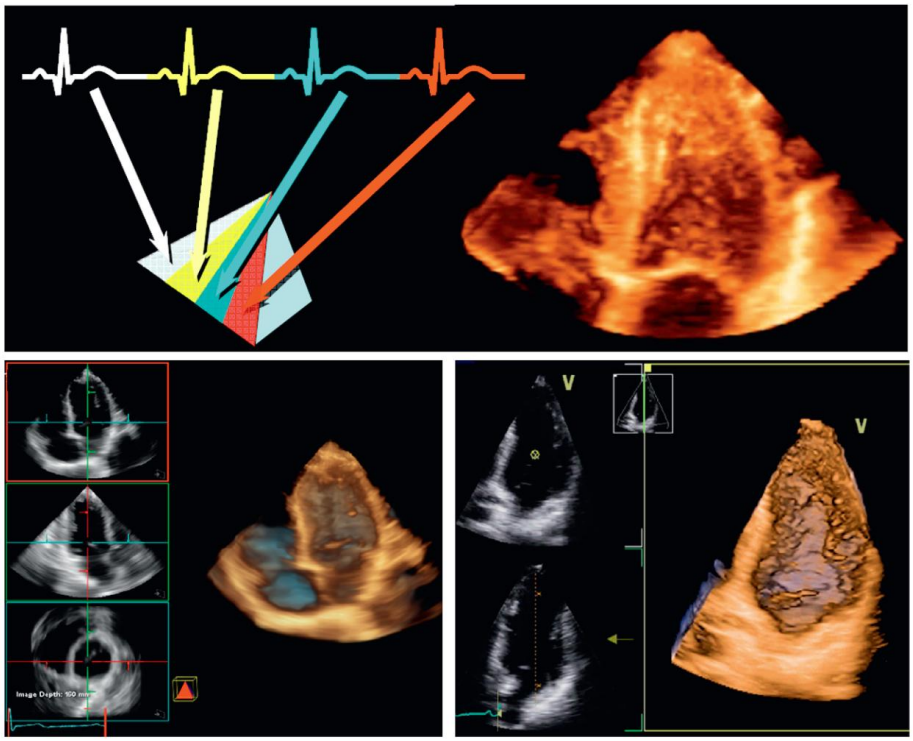


Figure 1-4 Transition from 'full-volume' acquisition relying on ECG-gated imaging of several subvolumes over a number of cardiac cycles (top) to single-beat acquisition mode (lower panels) (8)

Recently, significant advances in ultrasound, electronic and computer technology have thrust the field forward toward the development of a fully sampled matrix array transducer (7). Technological developments resulted in the capability to capture the entire heart in a single cardiac cycle (Figure 1-4, bottom panels). This approach promises to further improve the ease of RT3DE evaluation of the LV by improving the speed of acquisition and reducing artifacts (8).

1.2. Importance of cardiac chamber volume measurements

1.2.1. Parameters related with cardiac function assessments

Accurate, reproducible, noninvasive determination of cardiac chamber volumes, especially the left ventricular (LV) volumes, is important for clinical assessment, risk stratification, selection of therapy, and serial monitoring of patients with cardiovascular disease (10 to 12, 15 to 21). There are several parameters related with cardiac chamber volumes that are routinely assessed in clinical echocardiographic examinations: end diastolic volume (EDV), end systolic volume (ESV), stroke volume (SV), ejection fraction (EF), cardiac output (CO) and cardiac index (CI).

In cardiovascular physiology, EDV is the volume of blood in the ventricle at the end of filling (diastole), and it is used as a measure of cardiac global diastolic function. ESV is the volume of blood in the ventricle at the end of the cardiac ejection period (systole) and immediately preceding ventricular relaxation. It is used as a measure of systolic function. If the EDV and ESV are measured, then

$$SV = EDV - ESV;$$

$$EF = SV/EDV;$$

$$CO = SV \times \text{heart rate};$$

$$CI = CO/\text{body surface area (usually height}^2 \text{ in practice)}.$$

Those parameters are important predictors of cardiac morbidity and mortality. This will be demonstrated in the following two subsections, which takes two very common cardiac diseases, acute myocardial infarction and heart failure, as examples.

1.2.2. Acute myocardial infarction

Acute myocardial infarction (AMI) occurs when the blood supply to part of the heart is interrupted. It is mostly commonly (but not always) caused by occlusion of a coronary artery. This often follows the rupture of a vulnerable atherosclerotic plaque, which is an unstable collection of lipids and white blood cells on the wall of an artery. The resulting ischaemia and oxygen shortage, if left untreated for a sufficient period, can trigger a process called the ischaemic cascade: the heart cells die chiefly through necrosis and do not grow back. A collagen scar forms in its place. As a result, the patient's heart muscle tissue (myocardium) will be permanently damaged (10).

AMI continues to be a significant public health problem in developed countries and an increasingly significant problem in developing countries (9). The estimated mortality, although declining, still remains high. The prognosis in survivors of acute myocardial infarction depends on multiple factors, which relate both to the acute event, such as infarct size, location, and transmurality (whether the infarction is transmural or non-transmural), as well as to characteristics preceding the infarction such as age, sex, risk factors and prior infarction. However, the single most important determinant of survival is the functional status of the left ventricle after infarction.

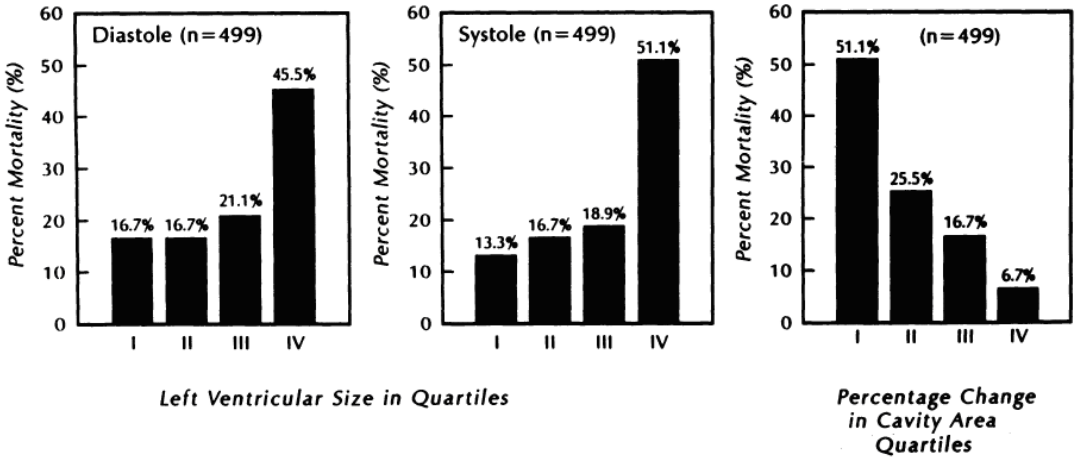


Figure 1-5 Relations between LVED area (left), LVES area (middle), and percentage change in cavity area (right) and subsequent cardiovascular mortality (10)

Figure 1-5 was obtained from two-dimensional echocardiographic quantitative measurements in 1994 (10). Short and long axis 2D transthoracic echocardiograms were obtained in 512 patients at a mean of 11.1±3.2 days after infarction and were repeated at

1 year in 420 survivors. The left ventricular size was assessed as left ventricular cavity areas (firstly calculate the two average areas of all short axis and long axis images respectively. Then the left ventricular size = average short axis area + average long axis area) at end diastole and end systole. The left ventricular contractile function was calculated as percent change in cavity area from end diastole to end systole. The 512 patients were divided into quartiles of increasing left ventricular areas of end diastole and end systole (left and middle graphs, I through IV). Both diastolic and systolic graphs demonstrate that an increase in area is associated with an increase in cardiovascular mortality. Percent change in cavity area (right graph, increasing from I through IV) shows that smaller percent change in area is strongly associated with higher cardiovascular mortality.

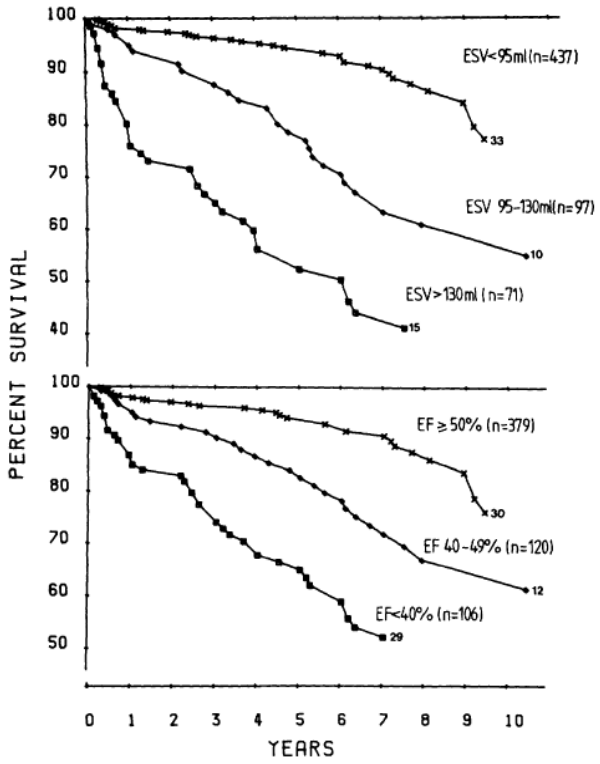


Figure 1-6 Actuarial survival curves constructed by dividing the patients into 3 groups according to their ESVs (top) or EFs (botton) (12)

In 1987, White et al (12) compared the individual predictive powers of LVEF, LVESV and EDV by measuring those parameters in 605 patients under 60 years old at 1 to 2 months after a first (n = 443) or recurrent (n=162) myocardial infarction. These patients were followed for a mean of 78 months for survivors (range 15 to 165 months).

One of the discoveries is shown in Figure 1-6 and Figure 1-7.

Figure 1-6 demonstrates that "although both methods of classification (LVESV or LVEF) give significant prediction for cardiac mortality ($p<0.001$), the separation is wider with classification by ESV than with classification by EF". Figure 1-7 shows that "Predictive value for ESV is apparent only when EF is less than 50%...Mortality for patients with EFs of 40% to 49% but ESV below the median being no worse than that for patients with EFs of 50% and above…". Based on these data, White and his colleagues concluded that "for prediction, end-systolic volume is the primary predictor of survival after myocardial infarction, being superior to ejection fraction when ejection fraction is low (<50%) or when end-systolic volume is high (>100ml)".

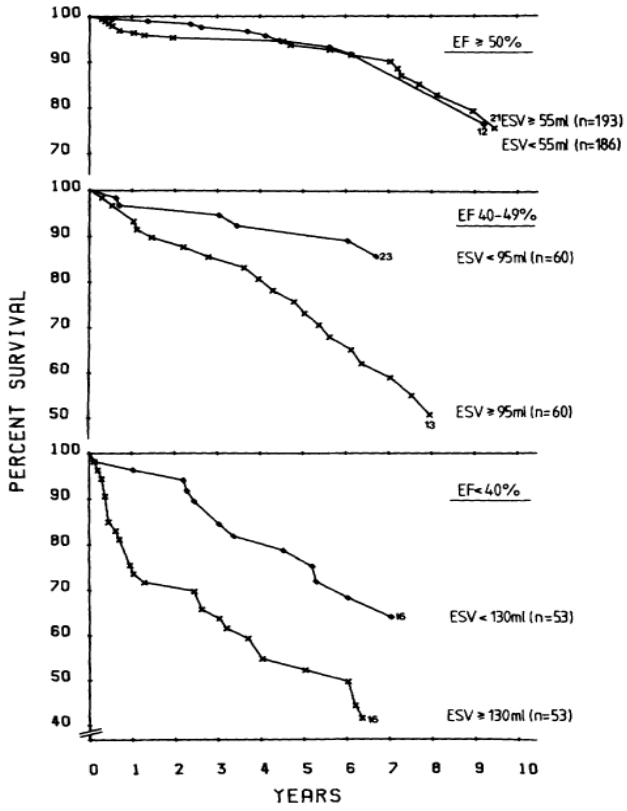


Figure 1-7 Actuarial curves constructed for 3 groups of EF, each group being subdivided according to whether ESV was above or below the median for that group (12)

Despite the differences and limitations of those studies, there is one obvious finding: the major predictor of long-term survival after recovery from acute myocardial infarction is the functional status of the left ventricle. Furthermore, National Institute for

Health and Clinical Excellence (NICE) clinical guideline and National Service Framework (NSF) for Coronary Heart Disease (CHD) indicate that left ventricular function assessment result influences the selection of treatment programmes, such as drug therapy, coronary revascularisation and selected patient subgroups (with or without left ventricular systolic dysfunction) after AMI (13, 14).

To better investigate the functional status of the left ventricle, accurate, non-invasive quantitative measurements of the ventricular volume are required. Therefore, imaging segmentation, edge detection and contour delineation techniques are designed and applied to quantitative volume measurements.

1.2.3. Heart failure

Heart failure (HF) is classically described as left ventricular dysfunction leading to congestion and reduced systemic perfusion, most often manifesting symptomatically as dyspnea and fatigue (15, 16). In general, HF is a cardiac condition that occurs when a problem with the structure or function of the heart impairs its ability to supply sufficient blood flow to meet the body's needs. In the United States, an estimated 5 million people have HF, and their ranks are increased by an estimated 550,000 each year (17). HF hospital stays have increased 150% over the last 20 years (17). It is the leading cause of hospitalization in people older than 65 (18). In Europe, 14 million people have HF and 3.6 million people are diagnosed with heart failure every year. 4.7% of hospitalizations in women, and 5.1% in men were due to HF (19). In some developing countries in Asia, the population of HF patient is also increasing due to economic development, industrialisation and urbanisation (20). With appropriate therapy, HF can be managed in the majority of patients, but it is a potentially life threatening condition, and progressive disease is associated with an annual mortality of 10% (21).

HF is caused by any condition which reduces the efficiency of the myocardium through damage or overloading, for example, myocardial infarction (in which the heart muscle dies of insufficient oxygen supply), hypertension (which requires extra force of contraction to pump blood), amyloidosis (in which heart muscle gets stiffer as protein is deposited in it), viral infection in children and alcoholism in adults, etc. After damaging the myocardium, the LV progressively dilates or hypertrophies. These morphologic

changes lead to further stress on the myocardium and increase the ventricular wall tension, which cause or exacerbate mitral valve regurgitation. All of these, in turn, results in further dilatation and hypertrophy in a vicious cycle (22). The whole process, which is called spherical remodeling, is often the final common pathway of HF development (Note: besides morphologic changes, there are other mechanisms and progressions of HF, for example, stiff heart muscle because of amyloidosis, and HF with preserved systolic function because of hypertension or cardiac ischaemia, etc).

HF can cause a large variety of symptoms, but many patients can be completely symptom free in the early stage. Because the morphologic process begins before the onset of symptoms, the recent HF guidelines place special emphasis on detecting sub clinical LV systolic and diastolic dysfunction. Although the majority of HF management decisions are guided by patient symptoms, precise and reproducible measurements of LV volume and EF are becoming more important in guiding important interventions (16). Consequently, quantified, objective measurements of LVESV, LVEDV and EF have become standard practice in echocardiography, which is commonly used to support a clinical diagnosis of HF, in many medical centers all around the world (16).

Several volume calculation methods have been applied to EF quantification in echocardiography, such as the ellipsoid method, summation of discs method and area-length method, for single and/or biplane calculation, respectively (For the volume quantification methods, more details will be given in Chapter 1.3). In general, abnormality is suggested as EF < 54%, and the cutoff-values for moderately abnormal and severely abnormal are 44% and 30% (23).

Besides the diagnostic purpose, EF and ventricular volumes have also been shown to provide crucial prognostic information in HF patients. Although EF and LV volume do not correlate with HF symptoms, they do provide significant correlations with the survival percentage of HF patients long-term. Morbidity and mortality are closely linked to both EF and LV volumes in multi-centre trials (16, 24, 25). Figure 1-8 is taken from Vasan et al's study in 1999 (24). In this study, they evaluated the echocardiograms of 73 subjects with CHF (33 women, 40 men, mean age 73 years) and 146 age- and gender-matched control subjects. 50% is defined as the cutoff between normal and

reduced LVEF (HF with LV diastolic dysfunction presents normal LVEF). From this figure we find that the overall survival of CHF subjects with reduced LVEF is worse than that of CHF subjects who have a normal LVEF.

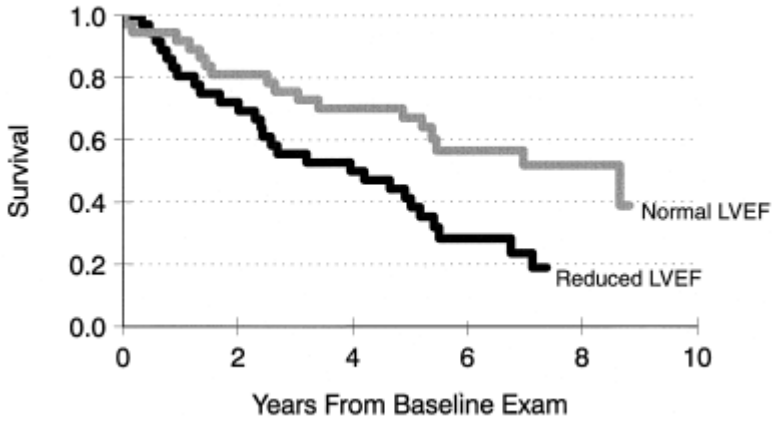


Figure 1-8 Kaplan-Meier survival plots of CHF patients with normal and reduced LVEF

1.2.4. Others

Besides AMI and HF, the accurate and reliable quantification of cardiac chamber volumes, especially the LV volumes and EF, also correlate well with prognosis and mortality across a wide range of cardiac diseases. The LVEF is also a frequent criterion upon which the decision to employ or withhold certain therapies depends, such as ICD (implantable cardioverter defibrillator) implantation, valve replacement, and coronary artery bypass surgery (13, 26 to 28).

1.3. Quantification of cardiac chamber volumes in RT2DE

RT2DE is the most widely used imaging modality for clinical assessment of various cardiovascular diseases. In clinical diagnosis, the LV volumes, the left and right atrial volumes can be calculated by different methods in RT2DE. This section will demonstrate what they are and discuss what drawbacks they have.

1.3.1. Clinical left ventricular volume calculation methods

Generally, in clinical 2D echocardiographic examinations, there are 3 methods that are

widely used to measure LVV and LVEF. Each of them is based on different volume calculation principles. They are the ellipsoid method, the area-length method, and the disc summation method (Simpson's rule) (Figure 1-9). According to the symmetric assumption, each of these 3 methods can be calculated by mono-plane or bi-plane formula. The apical 4 chamber and apical 2 chamber views are utilised for bi-plane calculation as these two views are nearly orthogonal (60 to 90 degrees) to each other. The three methods are demonstrated as follows:

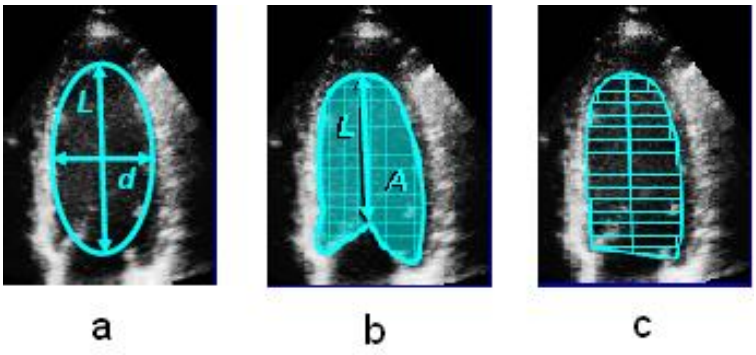


Figure 1-9 Three methods of measuring LV volume in 2DE (apical 2 chamber view): (a) Ellipsoid (b) Area-length (c) Disc Summation

(a) Ellipsoid method: left ventricular volume is approximated by an ellipsoid with diameter "d" and length "L".

Mono-plane (symmetric rotation): $V = \frac{4\pi}{3} \cdot \left[\frac{L}{2} \cdot (\frac{d}{2})^2 \right] = \frac{\pi}{6} \cdot L \cdot d^2$

Bi-plane: $V = \frac{\pi}{6} \cdot L \cdot d_a \cdot d_b$

(b) Area-length method: similar to ellipsoid method, it also assumes that the LV cavity closely resembles an ellipsoid. The area "A" and the length "L" are measured (62).

Mono-plane: $\because$ for an ellipse: $A = \pi \cdot \frac{L}{2} \cdot \frac{d}{2} \Rightarrow d = \frac{4A}{\pi L}$

$$\therefore V = \frac{\pi}{6} \cdot L \cdot d^2 = \frac{\pi}{6} \cdot L \cdot (\frac{4A}{\pi L})^2 = \frac{8}{3\pi} \cdot \frac{A^2}{L}$$

Bi-plane: $V = \frac{8}{3\pi} \cdot \frac{A_a A_b}{L}$

(c) Disc summation method (Simpson's Rule): left ventricular volume is approximated by "n" cylinders with diameter "di" and length "L".

Mono-plane: $V = \sum_{i=1}^{i=n} \pi \cdot (\frac{d_i}{2})^2 \cdot \frac{L}{n} = \frac{\pi}{4} \cdot \frac{L}{n} \cdot \sum_{i=1}^{i=n} d_i^2$

Bi-plane: $V = \frac{\pi}{4} \cdot \frac{L}{n} \cdot \sum_{i=1}^{i=n} d_{ia} \cdot d_{ib}$

1.3.2. Clinical left and right atrial volume calculation methods

Left and right atrial volume measurements can be obtained at ventricular systole when the volumes of the atria are at their largest. The parasternal short axis view and the apical two chamber view are utilized for the calculations. As with the LV volume calculation, the mono-plane ellipsoid method, the mono-plane area-length method and disc summation method can also be applied in atrial volume calculation (Figure 1-10, the volume calculation equations are the same as the LV volume mono-plane equations).

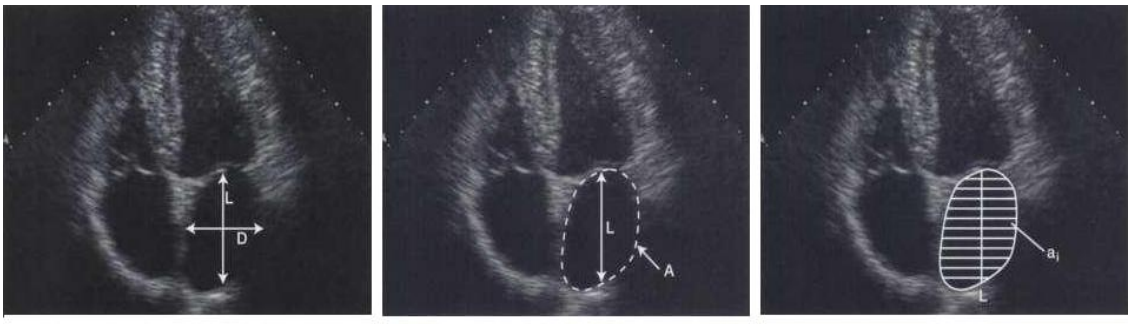


Figure 1-10 Three atrial volume calculation methods: a ellipsoid method; b area-length method; c disc summation method (29)

1.3.3. Clinical right ventricular volume calculation method

For right ventricular volume assessment, however, despite numerous attempts in 2DE, the techniques that have been developed so far are still very complicated. This is because of the irregular geometry of the right ventricle, the inability to visualize the

entire ventricle in any single 2DE plane and movement of the right ventricle with patient positioning. Therefore, the complexities of right ventricular volume measurements are greater than these of other cardiac chambers. Currently no techniques have been applied into the routine 2D echocardiographic examination for the calculation of right ventricular volume.

1.3.4. Disadvantages

Although RT2DE is widespread in clinical practice, it still holds significant drawbacks for volume measurements of heart cavities. These drawbacks cause inaccuracy as well as less reliability of heart cavity volume quantifications, such as measuring left ventricular volumes.

Geometric assumptions: the clinical techniques in RT2DE use symmetric and regular geometric models, such as a truncated ellipsoid, to approach the morphology of heart cavities, such as left ventricle. However, even healthy hearts do not conform well to these geometric models. Moreover, these models cannot take into account aneurysms, asymmetric ventricles or ventricular wall motion abnormalities (8, 30).

Foreshortening errors: foreshortening is difficult to avoid or even to recognize when acquiring apical long axis 2DE images (28), because there is no way to guarantee that a 2DE transducer is in the correct projection angle during imaging acquisition. Therefore, the foreshortening will cause underestimations of chamber volumes (30), as shown in Figure 1-11.

Tracing errors: the clinical measurement requires manual delineation of the cavity wall between the myocardium and the blood pool during the clinical 2DE assessment. The delineation is highly dependent on the quality of the 2D plane selected by the operator, and on the operator skill. Therefore, it is subjective and has intra- and inter-operator variability.

1.4. Quantification of heart left ventricular volumes in RT3DE

One of the most significant developments of the last decade in ultrasound imaging of heart was the evolution of three dimensional imaging from slow and labour intense

offline reconstruction to real time volumetric imaging. Real time 3D echocardiography (RT3DE) provides valuable clinical information that empowers echocardiographers with new levels of confidence in the diagnosis of heart disease (7, 30). This section will focus on the use of RT3DE in quantitative analysis of heart left ventricular volumes (LVEDV and LVESV) and function (EF). A series of aspects will be introduced, such as what advantages RT3DE can offer; what commercial software packages are frequently used for analyzing 3DE images; what imaging segmentation techniques (manual, semi-automated or automated) are developed for quantification of LV volumes and function.

1.4.1. Advantages of using RT3DE instead of RT2DE

The value of RT3DE imaging in the quantification of left ventricular volumes and ejection fraction has been demonstrated by multiple studies that compared RT3DE volume measurements with different reference techniques, including radionuclide ventriculography and cardiac magnetic resonance (CMR) (28, 31 to 58). These studies and others have demonstrated higher levels of agreement between the RT3DE approach and the respective reference technique compared with conventional 2DE methodology. Additionally, RT3DE measurements were found to be more reproducible than 2DE (28, 33 and 34) and, in some studies, even as reproducible as CMR (36) or even better than CMR (37).

The improvements in quantifying the LV volumes and function by using 3D imaging instead of 2D come from two sources: eliminating the need for very simple geometric modeling, which is inaccurate in the presence of aneurysms, asymmetrical ventricles, or wall motion abnormalities; eliminating errors caused by foreshortened views, as shown in Figure 1-11.

In Figure 1-11, for a same subject, the non-foreshortened RT3DE apical images (bottom panels) shows significant longer long-axis than the foreshortened RT2DE images (top panels). This explains that when comparing to the same reference technique, RT2DE showed bigger underestimation of the LV volumes than RT3DE.

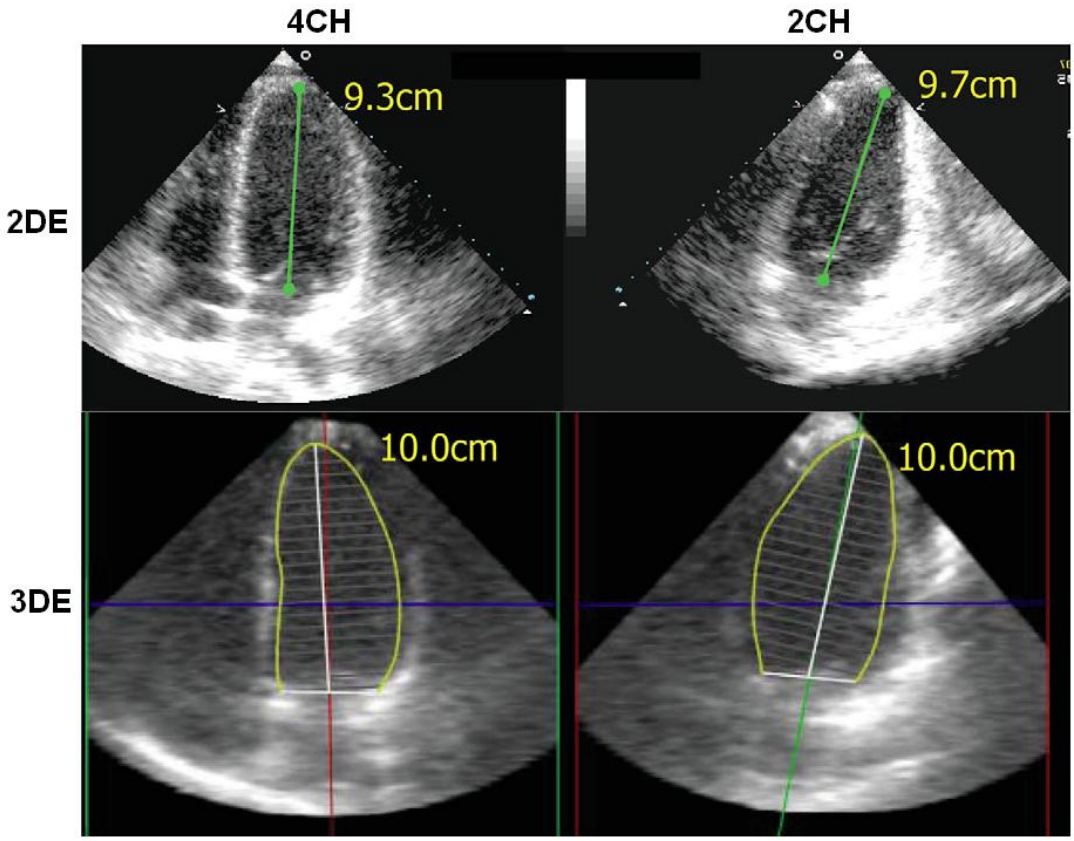


Figure 1-11 Effect of foreshortening on 2DE volume determinations and a method of using RT3DE data to reduce this source of error: the top panels represent 2DE apical 4-chamber (left) and 2-chamber (right) images with their respective long-axes indicated by the green lines. The bottom panels show the correct 4-chamber (left) and 2-chamber (right) images that have been extracted from the RT3DE dataset obtained in the same subject (28).

1.4.2. Commercially available volume measurement tools for RT3DE

Currently, for clinical practice, three commercially available software packages for RT3DE exist on the market: QLAB (Philips, Andover, Massachusetts, USA), TomTec 4D LV-Analysis (TomTec Imaging Systems, Unterschleissheim, Germany) and 4DLVQ (EchoPAC ver.108.1.0, GE Vingmed Ultrasound, Horten, Norway). All of them are frequently used programs for RT3DE analysis, but each of them uses different imaging analysis algorithms for measurement of LV volumes.

The similarities of all three tools are: they all use semi-automated methods to delineate the LV endocardial surface and calculate the LV volumes. Their workflows of volume quantification firstly require selections of proper apical images to avoid

foreshortening. Then, manual initializations are established for the automated border tracing. Finally, manual corrections are sometimes necessary if the programs traced incorrect borders. The differences are: TomTec requires triplane manual tracing of the endocardial border for initialization, whereas QLAB and 4DLVQ need input of several specific identification points (QLAB: five landmarks, 4DLVQ: nine landmarks). Therefore, the average time needed for measuring one LV volume and measuring accuracy as well as reproducibility is different for each tool (59 to 61).

1.4.3. Imaging segmentation techniques developed for LV volume quantification in 3DE

In the past decade, with the development of 3D echocardiography, efforts have been made for improved quantitative measurement of LV volume and ejection fraction. Measurement of the LV volume depends critically on LV surface delineation. Once the LV endocardial wall is detected, the volume calculation is very easy and fast. Several imaging segmentation algorithms have been developed for LV endocardial border delineation and volume calculation in 3DE, and some of them were commercialised. According to how much manual intervention is involved, they can be divided into manual and semi-automated and automated LV volume measurement techniques:

a. Manual: manually trace wall borders from short-axis planes + summation of discs (Simpson's rule): Simpson's rule is the basic principle of this technique. The LV is divided into several equidistant parallel short-axis slices spanning the LV cavity from the apex to mitral annulus (Figure 1-12 (A), in this figure there are 8 slices in total.). In contrast to the 2DE disc summation method, which only measures diameters of slices, and assumes each slice is an ideal cylinder, the real endocardial border is manually traced by hand-drawing closed contours in short-axis slices extracted from 3DE (as the white label in Figure 1-12 (B)). The volume of the traced region on each slice is calculated according to the number of pixels involved and the pixel dimensions. Adding up the volumes of all slices provides the LV volume at a chosen ventricular phase (41). The reconstructed LV cavity is shown in the right hand side of Figure 1-12 (A). This method is relatively simply but laborious and time consuming. Some publications have compared this method in 3DE volume calculation to the same or other techniques in

MRI, 2DE and radionuclide angiography (42 to 44). The results indicate that it shows excellent correlation with other methods in different imaging modalities.

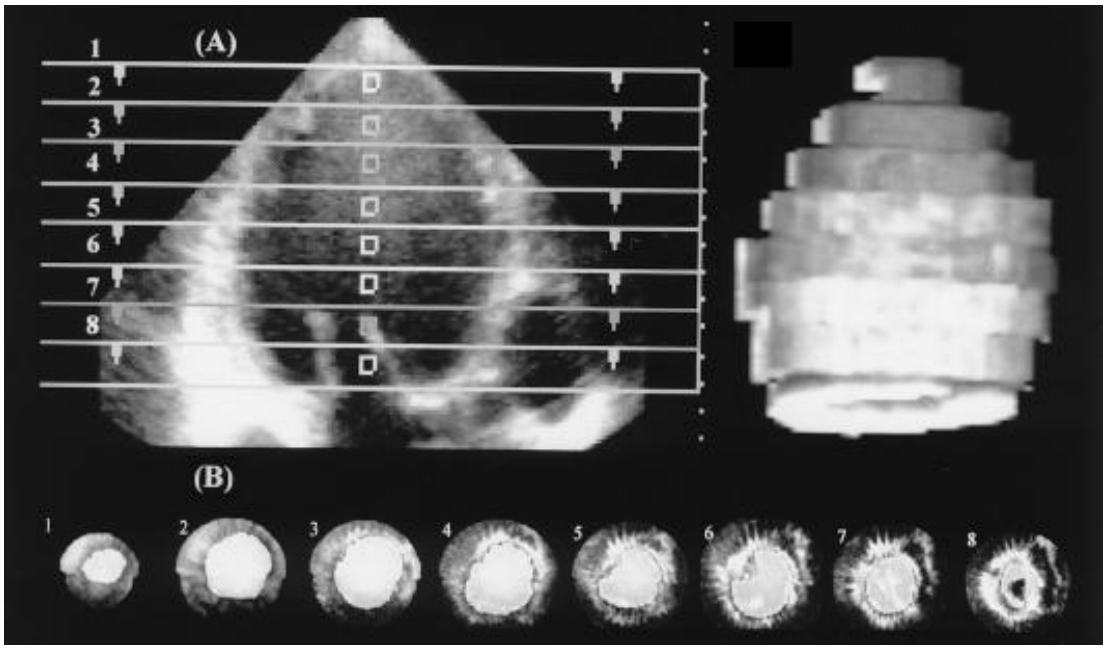


Figure 1-12 Principle of disc summation method in LV volume measurements (41)

b. Manual: manually trace wall borders from long- and short-axis planes + deformable shell model: Similar to the previous method, the endocardial borders are hand-traced in many 2D planes. However, the manual tracing was performed in both long- and short-axis planes this time. Then a minimum-energy deformable shell model is used to fit a closed elastic surface to the traced 3D endocardial borders (46). The surface fitting approach is based on the finite element method and parametric solid modelling (63). A 3D elastic surface is iterated towards hand-traced points from the endocardial surface (Figure 1-13) by varying different parameters of a set of functions that define the surface. This is known as “cost function” for optimization (38, 46 to 48, 63). There are no unique parametric functions of 3D surface model and model fitting, but the computation is made tractable by applying 2 physiological constraints: the surface is assumed to be locally smooth (no step edges on the ventricular surface), and the requirement that any point on the final surface be visible from the centroid of the surface (surface does not fold over on itself) (38, 46). After the surface fitting, volume is computed by summation of tiny tetrahedral elements formed by the vertices of a tessellation (gridding) of the surface and the centroid of the ventricular cavity.

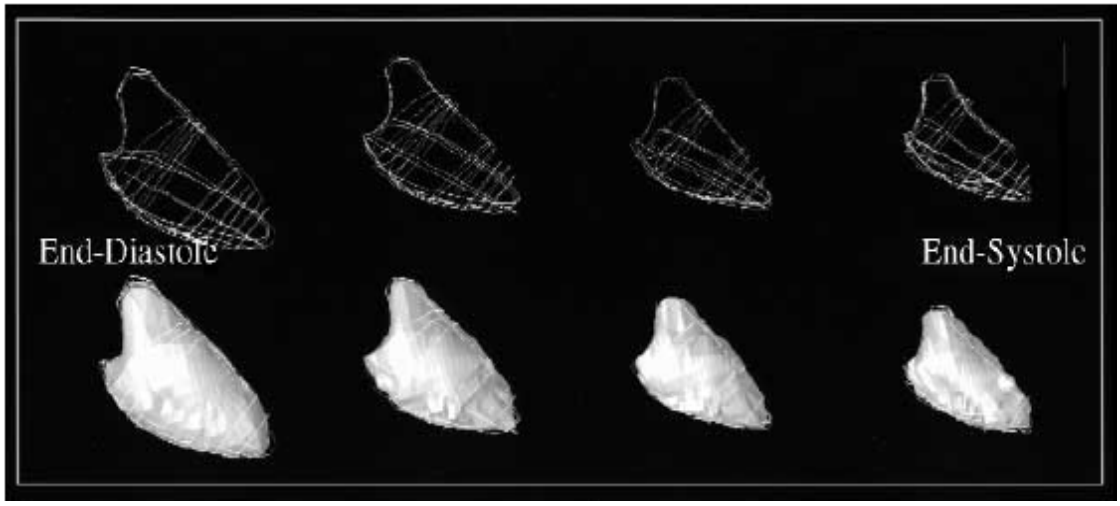


Figure 1-13 Manually traced borders (upper row) and the deformable shell approach (lower row) of LV from end diastolic to end systole in cardiac circle (46)

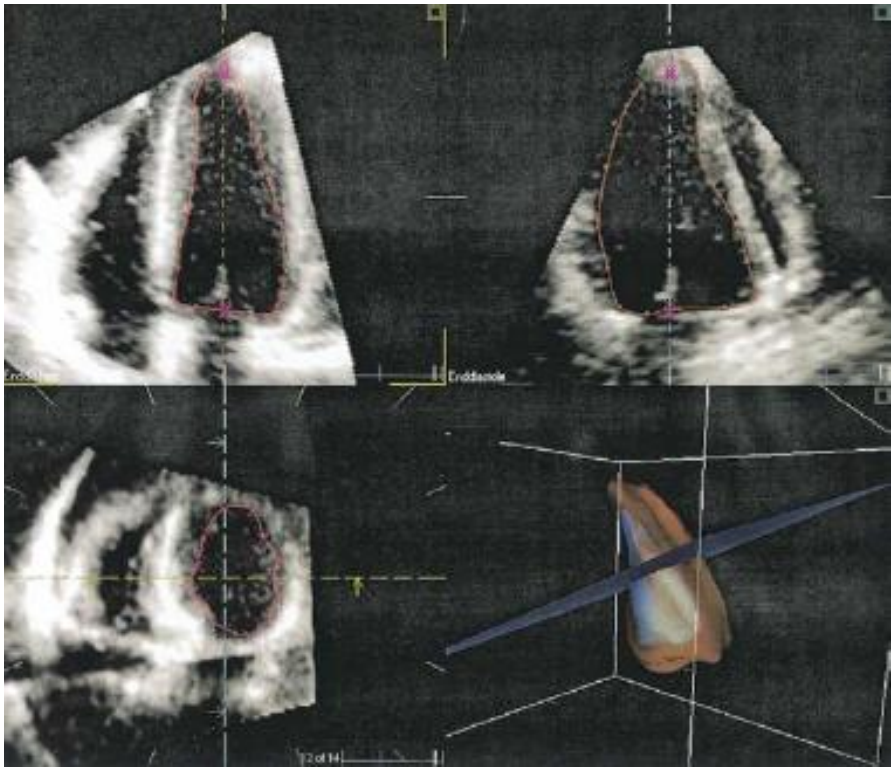


Figure 1-14 Tracing of the endocardial border in three planes and 3D reconstruction of the LV cavity (45)

c. Semi-automated: Tri-plane trace wall borders from long- and short-axis views + 3D reconstruction: The endocardial borders are manually traced by selecting some points on the edges in three different 2D planes including both short- and long-axis views. These points are then linked by curves to see whether extra adjustment is

necessary. Meanwhile the traced points are registered in a 3D space to reconstruct the LV cavity by a polyhedral surface reconstruction algorithm (45, 49). The basic idea of this algorithm is: connect the adjacent traced points to form hundreds of small triangles, so that hundreds of small surfaces are defined to build up a rough 3D surface of LV cavity. These small surfaces are then smoothed by decomposing them into even smaller surfaces. Finally, a 3D LV cavity is constructed, and the LV volume is calculated based on this model. Figure 1-14 demonstrates the tracing of the endocardial border in long-axis apical 2- and 4-chamber view (first row), and in short-axis view (second row, left) in the end-diastole phase of the cardiac circle. From those delineations, the 3D cavity of the LV is reconstructed (second row, right). The commercial software TomTec uses this technique (60).

d. Semi-automated: manual initialization on eight long-axis planes + spline model: Eight rotated long-axis 2D images are selected from the 3DE dataset (Figure 1-15 a) (50). In each of them, the LV apex and mitral annulus are identified. The markers at the mitral annulus are subsequently used to set the mitral valve plane and to truncate the contours of the LV cavity. Then, an ellipse is placed in each image, and adapted manually to fit as closely to the endocardial border as possible for initiation of the spline model (Figure 1-15 b). Thereafter 8 manually defined ellipses are set, a smooth spline model based on thin plate splines is initialized (51). Perpendicular to the model's contours, the algorithm seeks those image points which most likely define the endocardium (Figure 1-15 c). This is done by finite impulse response and morphologic filtering of grey-scale profiles perpendicular to the current contour estimate. Filter response and distance to the current contour are used to yield a likelihood estimate for each detected image point. And then, the spline model adapts its surface to approach those points automatically by considering the derived likelihood measures (50). Finally, a 3D endocardial surface is obtained for one frame. After that, the entire surface model is discretized as a polyhedron, and many small tetrahedral elements are formed, similar to the previous method. The volume is calculated by summation of the small tetrahedrons (50).

Besides calculating the LV volume, a similar model was also applied in assessment of

LV mass in 3DE in other publications. The results were compared with LV mass determined by MRI and M-mode echocardiography (52, 53), which shows that measurements in 3DE and MRI correlate well, but M-mode echocardiography overestimates LV mass.

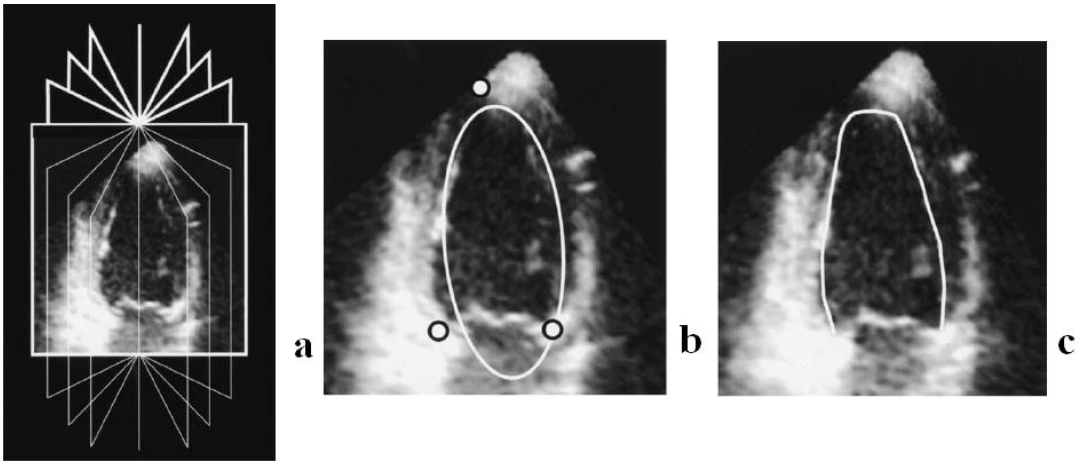


Figure 1-15 Eight planes selection and semi-automated delineation of the LV cavity based on spline model in 3DE (50)

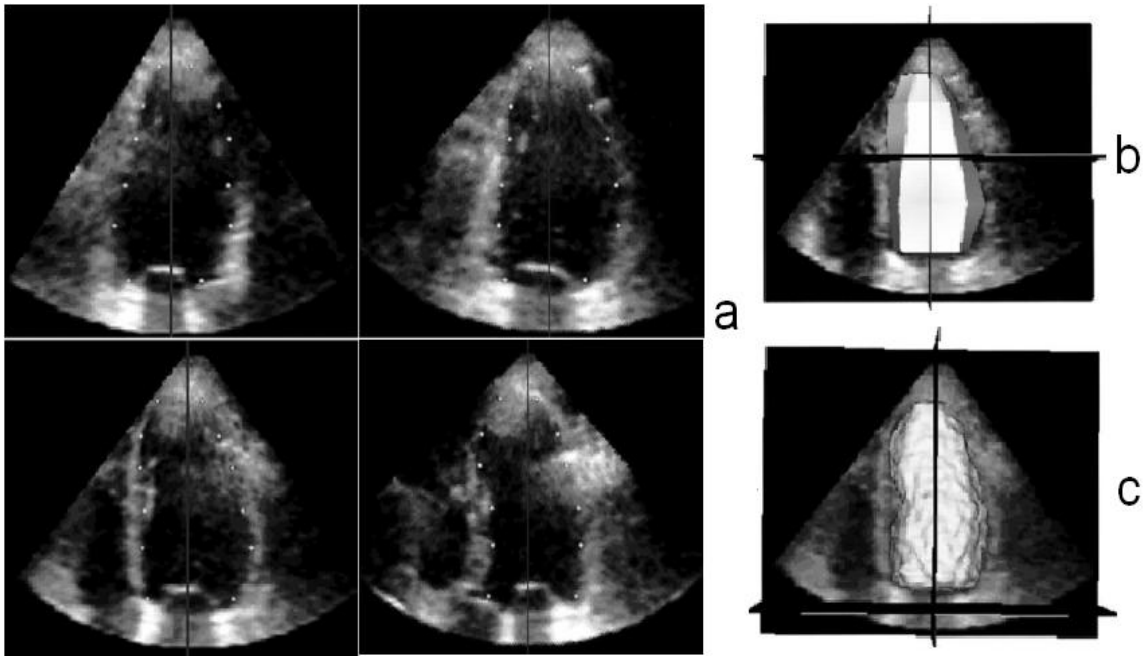


Figure 1-16 Four planes initialization and semi-automated detection of the LV cavity based on level set approach in 3DE (56)

e. Semi-automated: manual initialization on four long-axis planes + level set model: The LV surface detection is based on the level set approach (54), which uses an implicit representation of curves in the form of a partial differential equation to track

boundaries (55). Four planes rotated around the LV long axis at 45^{o} steps are selected from the 3D data set, and a small number of endocardial boundary points are manually initialized of each plane (Figure 1-16 a). The selected points are joined to define a rough surface, being the initial condition for level set partial differential equation (Figure 1-16 b), which guide the evolution of this surface within the volumetric dataset towards the endocardium under constraints of 2 forces: an interface tension force that depends on the curvature of the evolving surface and has a regulating effect, and a force that attracts this surface toward the image boundaries. When the 2 forces meet their balance after an iterative process, the evolution reaches a steady state, and the resultant surface is the detected endocardium (Figure 1-16 c). The LV volume can be measured by counting voxels confined within the surface (3, 56 and 58).

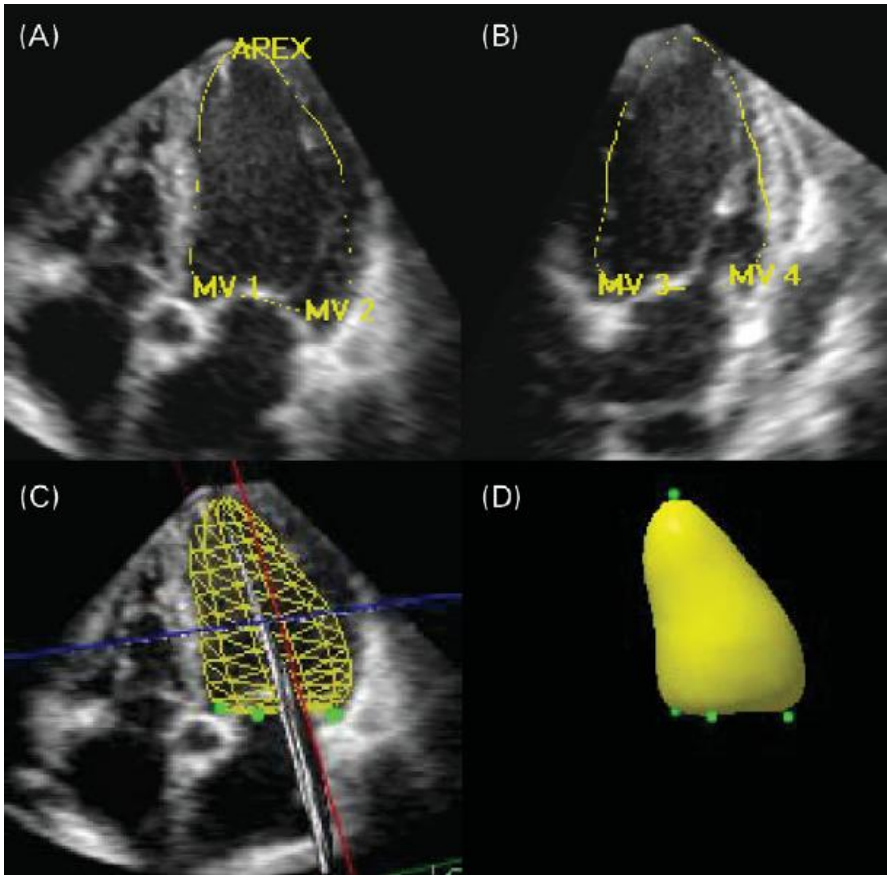


Figure 1-17 Five landmarks in 2 apical view and the resultant 3D LV endocardial surface from the deformable shell mode (28)

f. Nearly automated: manual initialization on two long-axis planes + deformable shell model: Only two planes are extracted from the 3D imaging dataset: apical 2- and 4-chamber views. Then five anatomic landmarks are manually initialized,

including 2 points to identify the mitral valve annulus in each of the two apical views and 1 point to identify the apex in either view (Figure 1-17 A and B: “MV1” to “MV4” and “APEX”). Following manual initialization of those points, the program automatically delineates the 3D endocardial surface using a deformable shell model (Figure 1-17 C and D) (57). If the 3D automatic delineation is not obtained perfectly, manual adjustments could be performed as well. This is particularly necessary in patients with LV aneurysms (The regional geometric distortion is too sharp to be adapted by the automated model). The LV volume is then automatically computed directly from voxel counts (28). Since identifying 5 points is the only manual intervention (most of the time manual correction is not necessary), this method is a nearly automated technique, and can measure the LV volume very quickly. The commercial software QLAB uses this technique (60).

1.4.4. Comparison between different methodologies in both 2D and 3D echocardiography

Despite many studies and publications on this topic, there is currently no perfect “gold standard” method for the determination of LV volumes using echocardiography. Table 1-1 lists published methods. Several aspects are given in the table: imaging modality (2DE or 3DE), type of intervention (manual, semi-automated or automated), number of subjects, “gold standards”, volume calculation time (Time 1), volume calculation time of each relevant standard (Time 2), correlation coefficient “r” and mean bias (Bland-Altman method) between methods and standards. The table clearly shows that the volume measurement techniques give higher correlation and smaller bias with their standards for 3DE in comparison with 2DE, which confirms that the accuracy of chamber volume measurement in 3DE is higher than it is in 2DE. The manually traced summation of discs method in MRI is recommended as gold standard in most publications. It normally requires more than 20 minutes calculating one volume. Some of the semi-automatic techniques in 3DE can be much faster. However, in fast semi-automatic techniques, the bias is larger than the bias in slow and manual traced techniques.

The limitation of clinical methods currently used in 2DE is mainly inaccuracy. This is

because, first, the foreshortening error is very difficult to avoid or even to recognize when acquiring apical long-axis view of the LV. The long-axis measured in 2D images is actually shorter than the real distance from apex to mitral annulus, which leads to an underestimate of the LV volume. Second, there is no real delineation of the endocardial surface in the clinical methods, but just an idealized geometric model. This assumption also cause serious inaccuracy since the diseased LV becomes asymmetric and differs from this model.

Although both of these shortcomings can be overcome by 3DE, and more accurate results have been confirmed, other drawbacks also appeared. Methods developed in 3DE are costly in terms of time and manpower. Because currently there is no automated way to delineate the endocardial surface, manual delineation or initialization is always required, which make them impractical for routine clinical use. Therefore, a trade-off between accuracy and operation time exists in quantitative LV volume measurements.

Despite the fact that 3DE is superior to 2DE in quantitative cardiac chamber volume measurements, it is still an echo based technique and suffers from the limitations inherent to the ultrasonic imaging modality. The spatial resolution is poor. The imaging contrast is week. The edge definition is not sharp. The noise and artifacts is serious. All of those affect the accuracy of quantitative measurements. Currently, MRI is widely used as the “gold standard” in LV volume assessments (Table 1-1).

1.5. Project aims

There were two aims of this research project. Firstly, it aimed to develop different new semi-automated algorithms for LV endocardial surface delineation, LV volume and EF quantification from clinical real time three-dimensional echocardiographic images. These algorithms were designed to have the following features:

- These algorithms required less manual initialization for LV endocardial surface tracing, so they were user friendly;
- These algorithms provided good accuracy in LV volume measurements.
- These algorithms provided good reproducibility in LV volume measurements.

Secondly, through assessing and comparing the performance of these newly developed algorithms in the aspects of accuracy and reproducibility, this project aimed to investigate what factors in real time 3D echo images influenced the performance of each algorithm, so that advantages and drawbacks of 3D echo images can be fully used and alleviated respectively.

Table 1-1 Different echocardiographic techniques for measurements of left venricular volume from publications

Technique	Methods	Modality	Intervention	Subjects	Standards	Time 1 (min/volume)	Time 2 (min/volume)	Correlation r	Mean Bias ± 2SD (Bland-Altman)	Reference
Clinical a	Bi-plane ellipsoid	2DE	Manual	10 healthy & 25 DCM	Bi-plane ellipsoid MRI	Not declared	Not declared	0.88	Not declared	Chuang ML et al. 2000 (38)
Clinical b	Bi-plane SOD	2DE	Manual	24 CAD, 18 DCM, 4 VD, 2 AI & 2 RAM	Manual trace SOD MRI	Not declared	Not declared	0.90	-19.0 ± 26.5 ml	Jacobs LD et al. 2006 (28)
Clinical b	Bi-plane SOD	2DE	Manual	4 Myocarditis, 9 HCM, 8 ICM, 6 VD, 2 CHD & 6DCM	Manual trace SOD MRI	Not declared	Not declared	0.82	-22.9 ± 119.3 ml	Gutiérrez-Chico JL et al. 2005 (45)
New a	Manually trace SOD	3DE	Manual	15 healthy & 31 ischaemic heart disease	Manual trace SOD MRI	15 to 18	30	0.98	-1.5 ± 24.0 ml	Nosir YFM et al. 1998 (41)
New b	Manually trace + DSM	3DE	Manual/Semi	8 healthy, 6 AF, 20 CAD, 10 DCM, 6 VD, 1 CHD & 4HCM	Manual trace ERA	14 to 24	Not declared	0.96	1.4 ± 10.3 %	Gopal AS et al. 1995 (49)

Technique	Methods	Modality	Intervention	Subjects	Standards	Time 1 (min/volume)	Time 2 (min/volume)	Correlation r	Mean Bias ± 2SD (Bland-Altman)	Reference
New b	Manually trace + DSM	3DE	Manual/Semi	12 healthy, 5 ICM & 8 DCM	Manual trace SOD MRI	Not declared	Not declared	0.99	-0.7 ± 11.3 ml	Hibberd MG et al. 2000 (46)
New b	Manually trace + DSM	3DE	Manual/Semi	10 healthy & 25 DCM	Manual trace SOD MRI	Not declared	Not declared	0.99	Not declared	Chuang ML et al. 2000 (38)
New c	Tri-planes reconstruction	3DE	Semiauto	9 HCM, 8 ICM, 6 VD, 2 CHD & 6DCM	Manual trace SOD MRI	Not declared	Not declared	0.99	-11.0 ± 29.1 ml	Gutiérrez-Chico JL et al. 2005 (45)
New d	8 planes spline model	3DE	Semiauto	10 healthy, 4 DCM & 10 ICM	Manual trace ellipsoid 3DE	12 ± 5	4 ± 1	0.99	Not declared	Kühl HP et al. 2004 (50)
New d	8 planes spline model	3DE	Semiauto	10 healthy, 4 DCM & 10 ICM	Manual trace SOD MRI	12 ± 5	14 ± 5	0.98	-16.7 ± 34.3 ml	Kühl HP et al. 2004 (50)
New e	4 planes level set model	3DE	Semiauto	11 healthy, 4 DCM & 6 AI	Manual trace SOD 3DE	Not declared	Not declared	0.99	-0.4 ± 5.2 ml	Corsi C et al. 2001 (3)

Technique	Methods	Modality	Intervention	Subjects	Standards	Time 1 (min/volume)	Time 2 (min/volume)	Correlation r	Mean Bias ± 2SD (Bland-Altman)	Reference
New e	4 planes level set model	3DE	Semiauto	6 healthy, 11 CAD, 9DCM & 2 VD	semiauto trace MRI	less than 15	40 to 60	0.98	-2.9 ± 20.0 ml	Corsi C et al. 2005 (58)
New f	5 points DSM	3DE	Close to Automated	24 CAD, 18 DCM, 4 VD, 2 AI & 2 RAM	Manual trace SOD MRI	2 to 7	Not declared	0.96	-10.2 ± 16.5 ml	Jacobs LD et al. 2006 (28)

Abbreviations and Acronyms: 2DE = 2-dimensional echocardiography; 3DE = 3-dimensional echocardiography; ERA = equilibrium radionuclide angiography; MRI = magnetic resonance imaging; SOD = summation of disks (Simpson's rule); DSM = deformable shell model; AF = atrial fibrillation; AI = aortic insufficiency; CAD = coronary artery disease; CHD = congenital heart disease; DCM = dilated cardiomyopathy; HCM = hypertropic cardiomyopathy; ICM = ischaemic cardiomyopathy; RAM = right atrial masses; VD = valvular disease.

Note: "Time 1" is the time required to generate one LV volume value in the echocardiographic method; "Time 2" is the time required to generate one LV volume in the corresponding standard. Both of them include the manual delineation/initialization time and the volume calculation time. The values in columns of "Correlation r" and the "Mean Bias±2SD" are average values of all phases in cardiac circle given in each publication.

Chapter 2. Algorithms

For accurate and reproducible quantification of left ventricular volumes, it is crucial to trace the endocardial wall surface on the echo images. Because of large variations of echo image quality and large variations of heart chamber geometry during dynamic cardiac cycles, the border delineation is normally performed manually or semi-automatically with manual initializations and interventions, as discussed in Chapter 1.

In this chapter, several newly developed semi-automated algorithms are introduced. They trace the endocardial wall borders in 3D echocardiography after a relatively small amount of manual setting. The procedures for establishing a region of interest, the functions of edge operators, the design of the border detection and interpolation algorithms and the mathematical principles of a finite element model are illustrated and explained one by one. This chapter specifically focuses on all the algorithms that were developed and used for performing different quantitative imaging analysis.

2.1. Selecting the region of interest (ROI)

A stack of 2D slices was acquired in the short axis view covering the entire left ventricle by 3D echocardiography (slice thickness: 1 voxel). Before applying the new algorithms, the region of interest (ROI) was selected manually as an initialization. This was the only procedure requiring manual interventions. There were two ways of ROI selection. Each of them designed for different analysis.

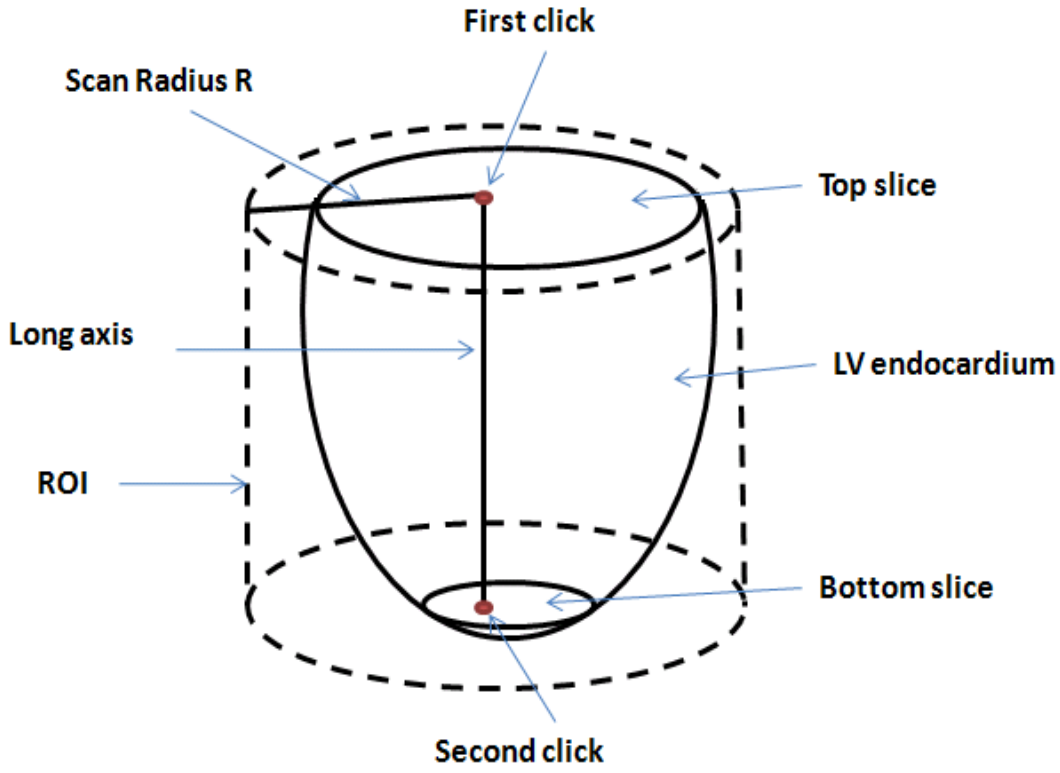


Figure 2-1 Manual initialization of the ROI, Method 1: after the first and second manual click, the ROI (dashed cylinder) was established.

Method 1: To establish the ROI, two short axis slices were extracted. One was at the top of the LV (LV mitral annulus); the other was at the bottom (close to LV apex). Then two points were selected, one on each slice by manual clicking inside the blood pool (Figure 2-1, first and second click). The basic principle for the operator to click the point was that it must be inside the blood pool and reasonably close to the blood pool centre. After that, a long axis coordinate was built up by connecting these two points, and together with the preset scan radius R, the ROI was established for all the LV short

axis slices automatically (Figure 2-1, the dashed cylinder). Note: the scan radius R must be set long enough to cover the whole blood pool areas for all the short axis slices.

Method 2: Different from Method 1, a third short axis slice located in the middle of the left ventricle was extracted together with the top and bottom slices. For each of the slices, two points were chosen to set the scan radius, one in the blood pool and the other in the myocardium. Therefore, there were six manual clicks, and three scan radii in total (Figure 2-2, first to sixth click, scan radius R1, R2 and R3). Then the ROI (dashed area) was established automatically as shown in Figure 2-2. The length of the scan radius R was changing linearly from R1 to R2 and from R2 to R3. Like Method 1, the click inside the blood pool should be reasonably close to the centre. The scan radius R must be set long enough to cover the whole blood pool areas for all the short axis slices.

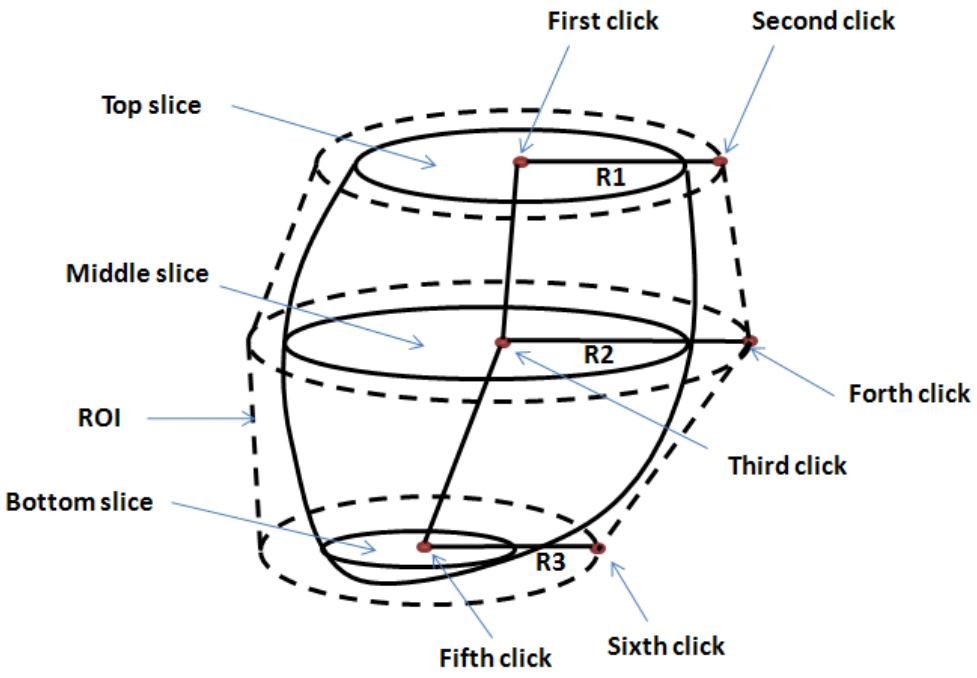


Figure 2-2 Manual initialization of the ROI, Method 2: after six manual clicks on three slices (Top, middle and Bottom), the ROI (dashed area) was established.

Compared to Method 1, Method 2 requires more manual initializations. This is necessary because the geometry of the left ventricle varies for every individual. Sometimes, Method 1 cannot establish a proper ROI for a LV with a very irregular shape. The ROI is like a cage that needs to wrap the whole endocardial surface of the LV inside. If any part of the endocardial surface is not contained inside the ROI, it

would be impossible for the algorithms to detect the correct endocardial border of that part. However, if too many irrelevant parts are contained inside the ROI, such as the epicardial surface or fat tissues around the myocardium, it would raise the possibility of detecting the endocardial borders incorrectly.

2.2. First order differential gradient operator

After establishing ROI, first order differential gradient operators (also called edge operators) were applied to echo images. The original intensity images were transferred to edge enhanced gradient images (Figure 2-3). In the algorithms, two different gradient operators were employed (coded as G_1 and G_2):

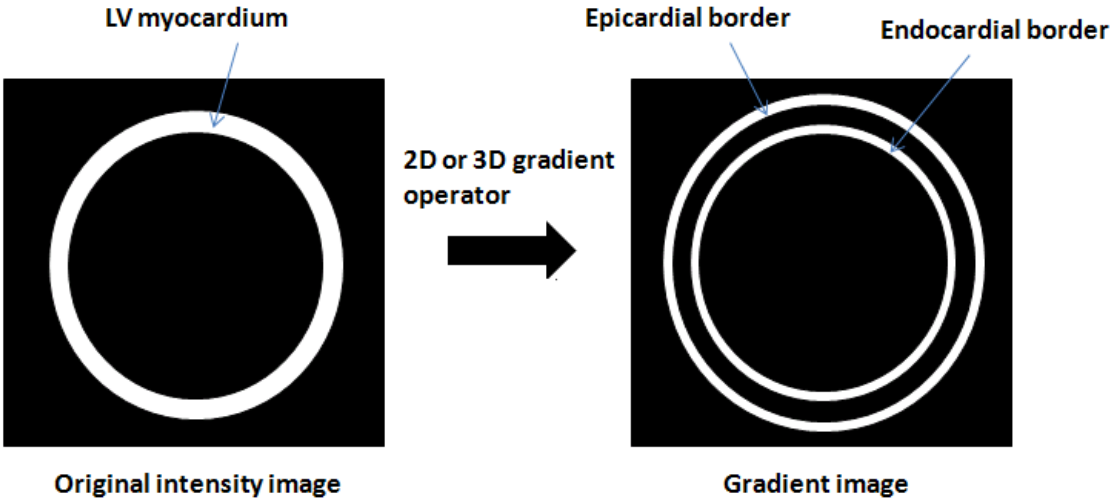


Figure 2-3 A demonstration of applying first order gradient operator on a simplified 2D short axis slice of left ventricle

2D gradient operator (G_1): All the 2D short axis slices were transferred to 2D gradient images one by one through the following convolution relationships (64):

$$G_x(j,k) = I(j,k) \otimes H_x(j,k) \tag{2.1}$$

$$G_y(j,k) = I(j,k) \otimes H_y(j,k) \tag{2.2}$$

then

$$G(j,k) = \left\{[G_x(j,k)]^2 + \left[G_y(j,k)\right]^2\right\}^{1/2} \tag{2.3}$$

where I(j, k) is a 2D slice extracted from a 3D image (j and k are the pixel indices in x- and y-dimension); H_x(j, k) and H_y(j, k) are n × n row and column impulse response arrays; G(j, k) is a 2D gradient image.

3D gradient operator (G_2): Instead of applying 2D gradient operator slice by slice, 3D echo image was directly convolved with 3D gradient operators as following:

$$G_x(j,k,l) = I(j,k,l) \otimes H_x(j,k,l) \tag{2.4}$$

$$G_y(j,k,l) = I(j,k,l) \otimes H_y(j,k,l) \tag{2.5}$$

$$G_z(j,k,l) = I(j,k,l) \otimes H_z(j,k,l) \tag{2.6}$$

then

$$G(j,k,l) = \left\{[G_x(j,k,l)]^2 + \left[G_y(j,k,l)\right]^2 + [G_z(j,k,l)]^2\right\}^{1/2} \tag{2.7}$$

where I(j, k, l) is a 3D image (j, k and l are the voxel indices in x-, y- and z-dimension); H_x(j, k, l), H_y(j, k, l) and H_z(j, k, l) are three n × n × n impulse response arrays in three orthogonal directions, and G(j, k, l) is a 3D gradient image. (Note: For both 2D and 3D gradient operators, n must be odd.)

For ultrasound images, the signal to noise ratio can be poor (4), and acoustic noise can confound the detection of cardiac chamber wall edges. This problem can be alleviated by extending the size of the neighbourhoods over which the differential gradients are computed. Argyle (77) and Macleod (78, 79) have proposed large neighbourhood Gaussian-shaped weighting functions as means of noise suppression, which defines our 2D and 3D gradient operators. Let

$$g(x,s) = [2\pi s^2]^{-1/2} exp\{-1/2(x/s)\}^2 \tag{2.8}$$

denote a continuous domain Gaussian function with standard deviation s (s is also called spread parameter which controls the width of the Gaussian curve). Utilizing this notation, the Macleod horizontal gradient impulse response function is given by

$$H_x(x,y) = [g(x+s,s) - g(x-s,s)]g(y,t) \tag{2.9}$$

where s and t are spread parameters in x and y dimensions. The vertical impulse response function H_y is the transpose of H_x.

The Macleod operators give decreasing importance to pixels (voxels) far removed from the centre of the neighbourhood. An example of 2D 5×5 normalised row and column gradient operators is:

$$H_x = \frac{1}{23}\begin{bmatrix} 1 & 2 & 0 & -2 & -1 \\ 2 & 3 & 0 & -3 & -2 \\ 3 & 4 & 0 & -4 & -3 \\ 2 & 3 & 0 & -3 & -2 \\ 1 & 2 & 0 & -2 & -1 \end{bmatrix}, H_y = \frac{1}{23}\begin{bmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 3 & 4 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -3 & -4 & -3 & -2 \\ -1 & -2 & -3 & -2 & -1 \end{bmatrix}.$$

There is a trade-off between noise reduction and resolving power. As the size of the gradient operator increases, the influence of the noise will be reduced and vice-versa, but the apparent width of the edge will be broadened. Therefore, an optimal size of the gradient operators exists. Furthermore, for different images, the spatial resolution may be different: the pixel (voxel) size may represent different real lengths. This also affects the choice of optimal operator size.

In general, 3D gradient operators make use of neighbourhood voxels around each individual pixel in 3D, rather than 2D operators which are only able to combine information in a plane. For example, apparent gaps in a vessel or chamber wall are very common in ultrasound images, which are normally caused by noise and artefacts. These gaps are missing information in one 2D plane, and it is impossible to restore them no matter how the 2D operator size is extended to gain more information from this plane. But they may be restored from the intensity information from neighbouring planes by utilizing 3D gradient operator. However, the noise in neighbouring planes can also affect the original one through the 3D operators, which cause a smoothing effect and lower the image contrast. This phenomenon is demonstrated in the next chapter.

2.3. Border detection algorithms

The border detection algorithm was designed to seek the endocardial wall borders from one point in the blood pool towards the endocardial wall surface along a set of radial lines (Figure 2-4).

After the initialization of the ROI, each 2D slice had a cross point with the ROI long axis (Figure 2-4, red point). From this cross point, the image pixel intensity value of both original slice and gradient slice along a set of radial lines was scanned with preset radius R (Figure 2-1) and angle Δθ (default: Δθ = 3°) in a counter-clockwise direction.

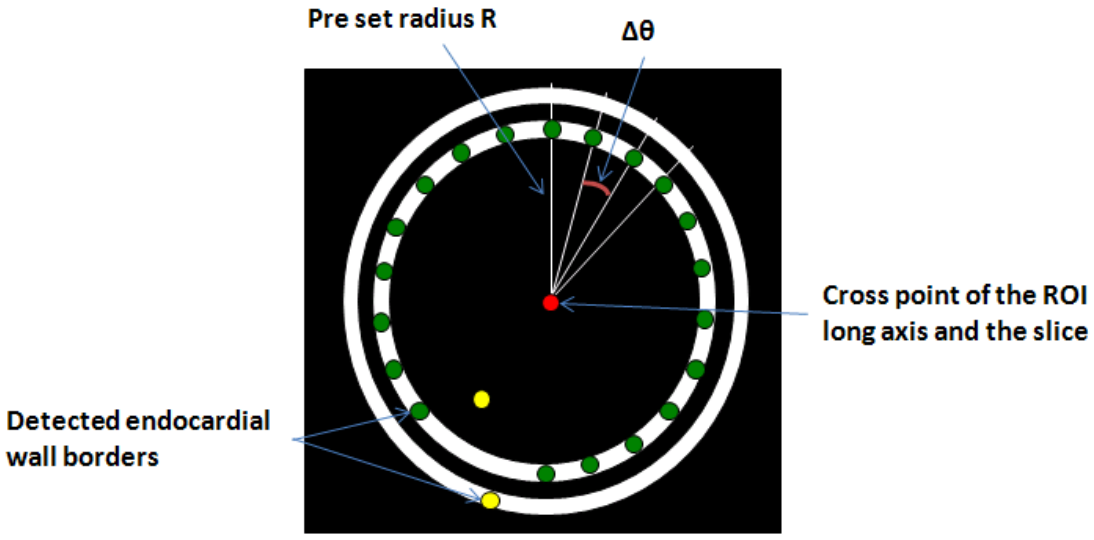


Figure 2-4 A demonstration of border detection principle: seeking borders from the cross point towards endocardial wall boundary along a set of radial lines: green points: true detection; yellow points: false detection

In a normal situation, for each radial line, the intensity value of the balloon wall pixels was larger than the value of the blood pool pixels (Figure 2-5, A), and the gradient operator generated two peaks (endo- and epicardial wall edge gradients, Figure 2-5, B). In order to detect the endocardial wall borders in the gradient images, a root mean square (rms) of each radial line was applied as a threshold:

$$T = c \cdot \sqrt{\frac{1}{n}\sum_{i=1}^{i=n} I_i^2} \qquad (2.10)$$

where T = threshold; c = coefficient (preset value can be 1.0 to 1.5; default value: 1.2); I_i = pixel intensity; n = number of pixels along a radial line. T is roughly proportional to the signal to noise ratio along each radial line (64), thus it was selected as a cut-off value. After applying the threshold, the plots of the gradient slice were transferred to binary plots, and only the two peaks (endo- and epicardial wall edge gradients) whose

magnitudes were higher than the threshold become “1”. The rest of the pixels were “0” (Figure 2-5, C). The endocardial wall boundary was then located on the first peak from the cross point.

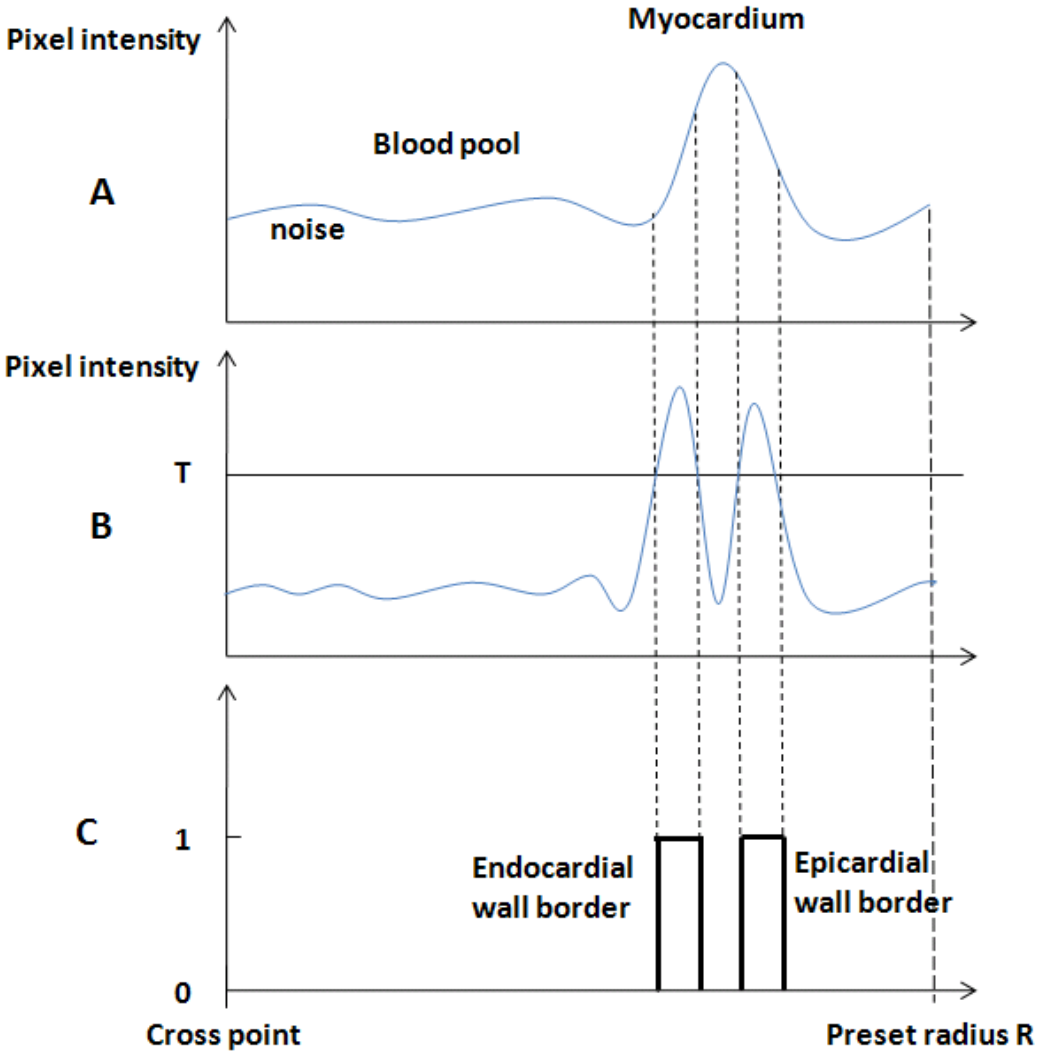


Figure 2-5 A sketch of Pixel Intensity vs. Scanned Radius R in a normal situation: A. intensity profile of original image; B. intensity profile of gradient image; C. binary plot obtained after applying threshold; Edges around endo- and epicardial wall border became “1”

To detect the endocardial wall border automatically, border detection algorithm 1 (coded as B_1 for short) was developed by analysing the pixel intensity profile in normal situations. The basic procedure is as following (Figure 2-6):

Border detection algorithm 1 (B_1): First detect the pixel with the largest intensity value in each scanned radial line of the original slice, which was the highest point in the

"Pixel Intensity vs. Scanned Radius R" plot of the original slice (Figure 2-6, step 1). This point was considered to be inside the myocardium. Then in the corresponding binary plot of the gradient slice (Figure 2-6, step 2), the program sought the first binary peak from this point towards the cross point (Figure 2-6, step 3). If it existed, the central pixel of the binary peak (Figure 2-6, step 3, the red dot) was considered to be the endocardial wall border, and the distance from the endocardial wall border to the cross point was saved as radii r(θ) ($\theta \in [0, 2\pi]$). If it did not exist, the detection along this scanned radial line was marked as failed, and r(θ) = 0.

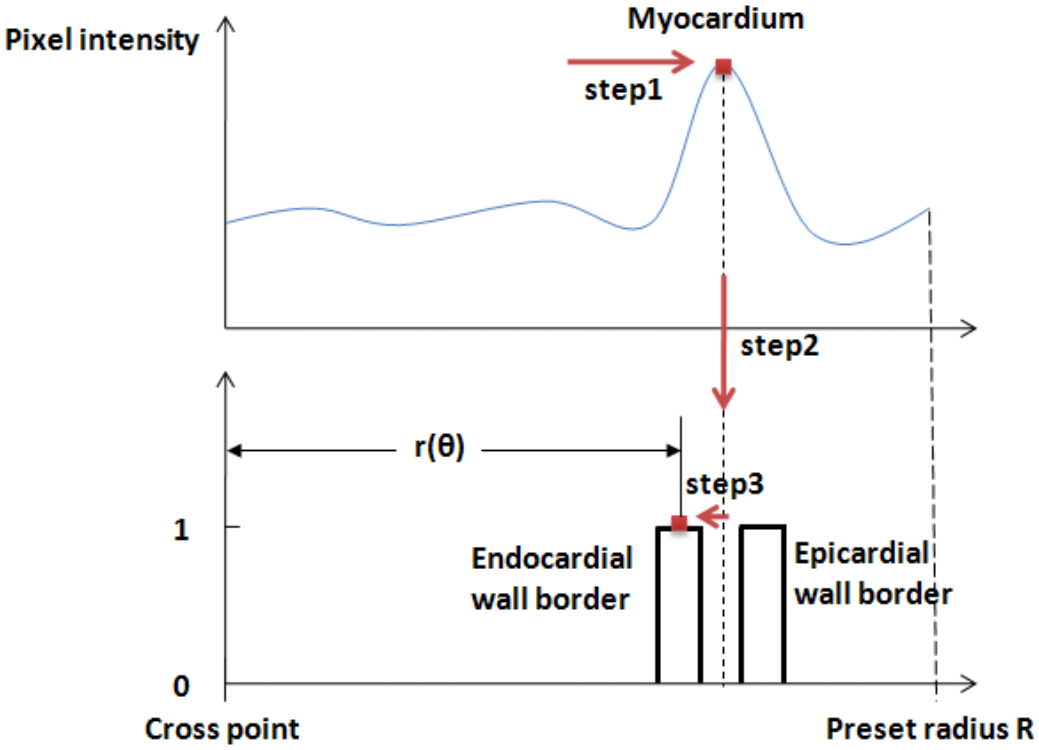


Figure 2-6 Basic procedure of border detection algorithm 1 (B_1) in the same situation as Figure 2-5: B_1 followed the step 1, 2 and 3 to find the border.

In practice, the B_1 was not always correct due to the complexity of the intensity variation caused by noise and artefacts in ultrasound images. For example, sometimes the noisy pixels inside or outside the left ventricle chamber instead of the myocardium pixels might have the largest intensity value (Figure 2-7). As a result, B_1 considered an incorrect location as the myocardium in step 1. Once this happened, it was impossible that the following step 2 and 3 could find the correct border of the endocardium. Therefore the second border detection algorithm was introduced:

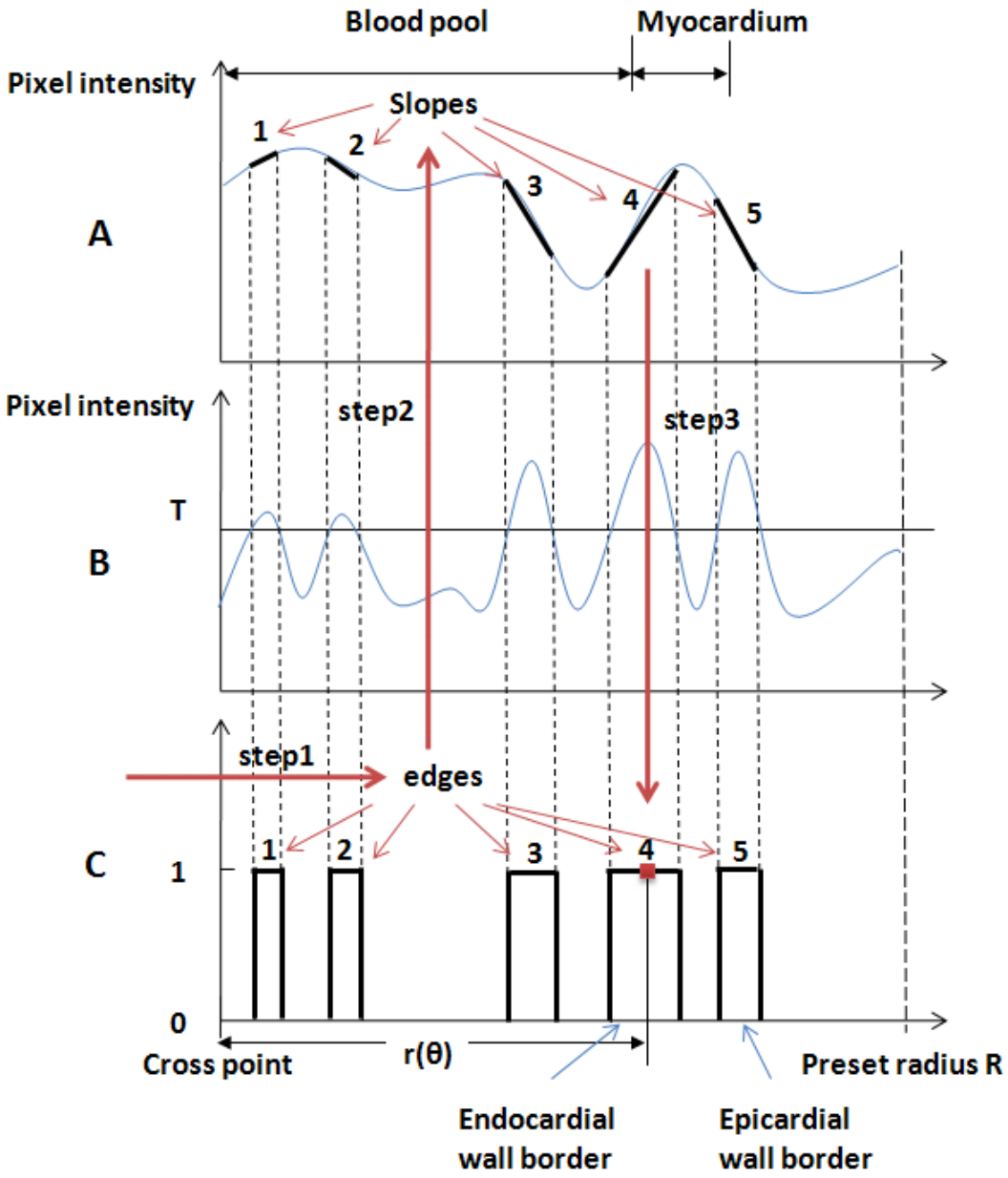


Figure 2-7 Basic procedure of border detection algorithm 2 (B_2) in a serious noisy situation: B_2 followed the step 1, 2 and 3 to find the border

Border detection algorithm 2 (B_2): Different from B_1, B_2 firstly located all edges (peaks with pixel values equal one) in the binary gradient plots (Figure 2-7, C. step 1: edges 1, 2, 3, 4 and 5). Then the algorithm went back to the original intensity profile of this radial line, and looked for the corresponding pixel intensities of these edges (Figure 2-7, step 2). After that, B_2 performed the following analysis procedure in step 3:

1) Calculated the linear slope value of each edge: Figure 2-7, A: slope 1, 2, 3, 4 and 5.

2) Eliminated all the negative slope values: Figure 2-7, A: slope 2, 3 and 5.

3) For positive slopes, considered the one with the largest slope value as the endocardial wall slope (Figure 2-7, A: slope 4).

4) Went back to the binary gradient plots, and found out the corresponding edge (step 3: Figure 2-7, C: edge 4). The distance from the central pixel of the edge (the red dot in edge 4) to the cross point was saved as radii r(θ) ($\theta \in [0, 2\pi]$): Figure 2-7, C, r(θ).

5) If there were two or more edges showing the same slope value, the detection along this scanned radial line was marked as failed, and r(θ) = 0.

6) If there were no edges showing positive slope value, the detection along this scanned radial line was marked as failed too. r(θ) = 0.

Both B_1 and B_2 have their advantages and disadvantages. B_1 is more sensitive, and requires less computational time. It normally finds the strongest edge along each scanned radial line, although some of them may not be the correct endocardial wall borders. For example, in the case of Figure 2-7, B1 is likely to consider the edge 1 as the endocardial wall border. B_2 is more specific. It involves all the detected edges in the gradient image, rather than just one edge. It investigates the intensity information of the original image, and makes comparisons in order to decide which edge is most likely to be the correct endocardial wall border. However, for some cases, it may consider none of the detected edges to be the endocardial wall border.

There are two reasons of setting r(θ) = 0 if the detection failed. One is that it can clearly indicate the border detection along which the radial line failed. The other is that it can let the volume calculation algorithm exclude the failed detection areas for calculating ventricular cavity volumes. This becomes clear in section 2.5.

2.4. Border interpolation algorithms

Neither B_1 nor B_2 could trace the endocardial wall borders perfectly. Because of the poor spatial resolution and high noise in echo images, as well as the complexity of the LV geometry, it was normal that the border detection algorithms could not detect the

border, or considered wrong pixels as the border along some of the scanned radial lines (Figure 2-4, yellow points). To overcome these problems, border interpolation algorithms were developed.

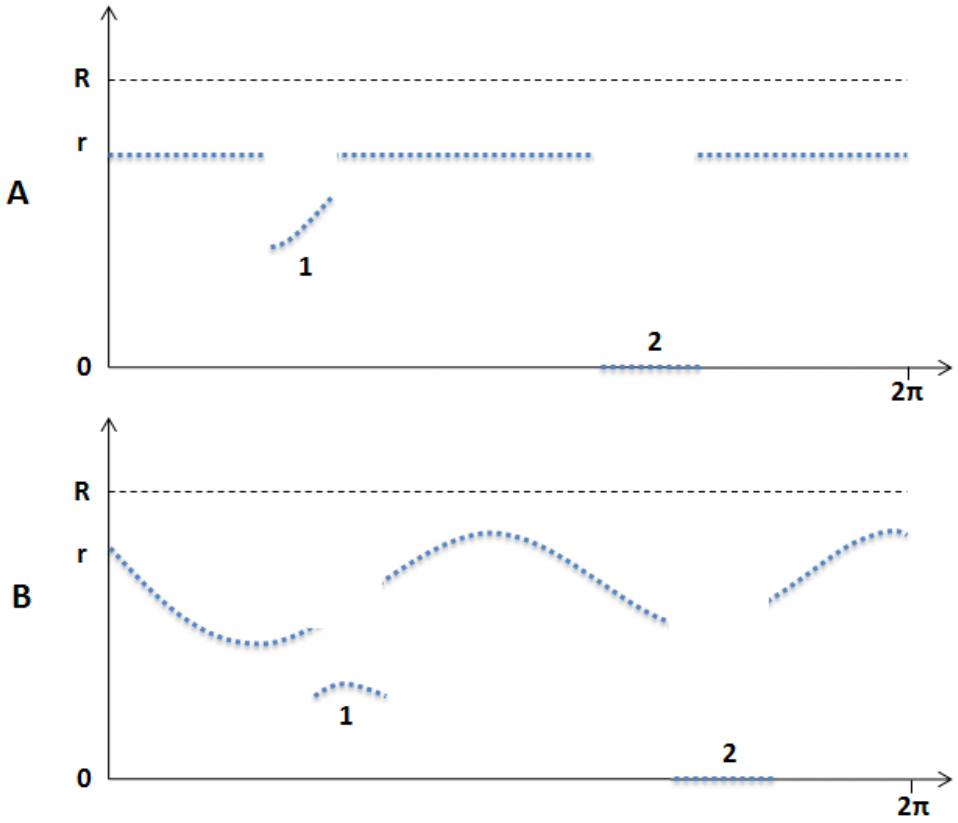


Figure 2-8 A demonstration of plot r(θ): A. 2D LV boundary in short axis view was a perfect circle. B. 2D LV boundary in short axis view was an ellipse liked shape. Note: in both cases, '1' represents the incorrect detection; '2' represents the failed detection.

Linear interpolation (I_1): After border detection algorithms, r(θ) were obtained, where $\theta \in [0, 2\pi]$ and r $\in [0, R]$. In a few cases, the LV boundary in 2D short axis view looked like a perfect circle. If r was plotted as a function of θ, a plot as Figure 2-8 A would be obtained, where r(θ) = r. However, in most cases, the LV boundary was an ellipse liked shape, rather than a perfect circle (Figure 2-8, B). Therefore, r(θ) was not a constant value. Figure 2-8 showed that in both A and B, border detection algorithms might have some incorrect detection and failed detection, such as '1' and '2'. To correct them, linear interpolation (I1) drew a straight line to connect across incorrect or failed detections, as shown in Figure 2-9, red lines.

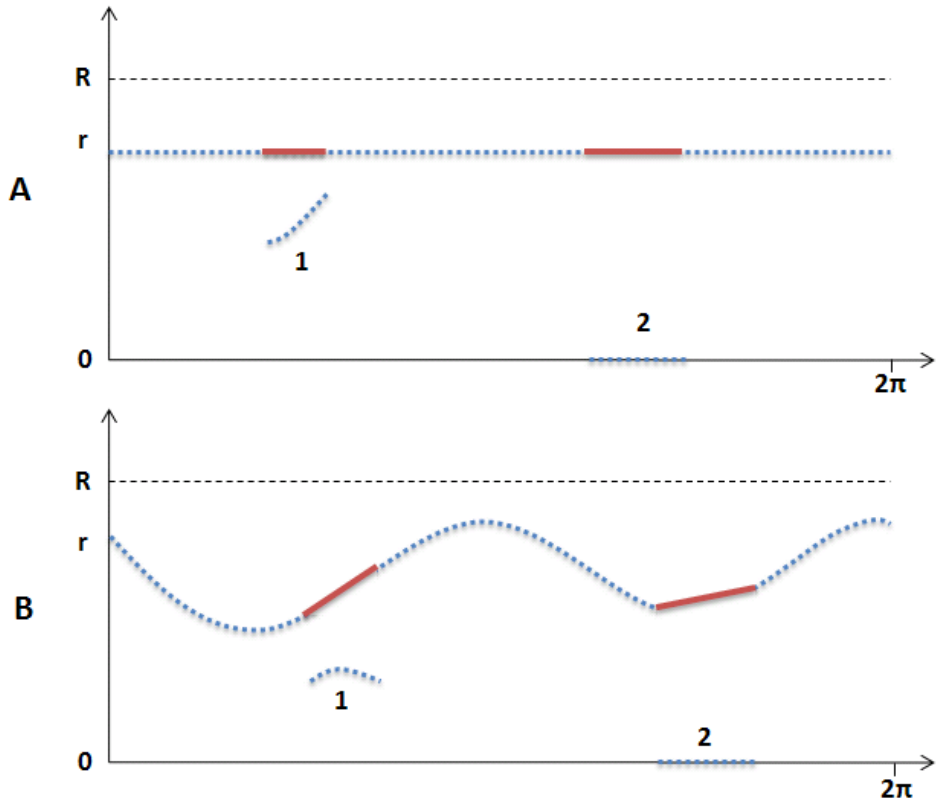


Figure 2-9 A demonstration of linear interpolation algorithm 1 (I_1): I_1 applied a straight line to connect across incorrect of failed detections, which is represented by the red lines in A or B.

Ellipse interpolation (I_2): I_2 assumed that the geometry of the LV in short axis view should be an ellipse (or a circle in some cases). In the Cartesian coordinate system, the general equation of an ellipse is:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \qquad (2.11)$$

with A, B, C not all zero, and $B^2 - 4AC < 0$. If A = C and B = 0, the equation represents a circle. After normalising *F*, it changes to:

$$ax^2 + bxy + cy^2 + dx + ey + 1 = 0 \qquad (2.12)$$

I_2 used the least square method to fit an ellipse into detected borders: by minimizing

$$\sum_{i=1}^{i=n} \|(x,y)_i - (X,Y)_i\|^2$$

$$X = rcos\theta$$

$$Y = rsin\theta$$

where $(x, y)_i$ is a point on ellipse; $(X, Y)_i$ is a detected endocardial wall border; n is the number of scanned radial lines, parameters a, b, c, d and e were approached.

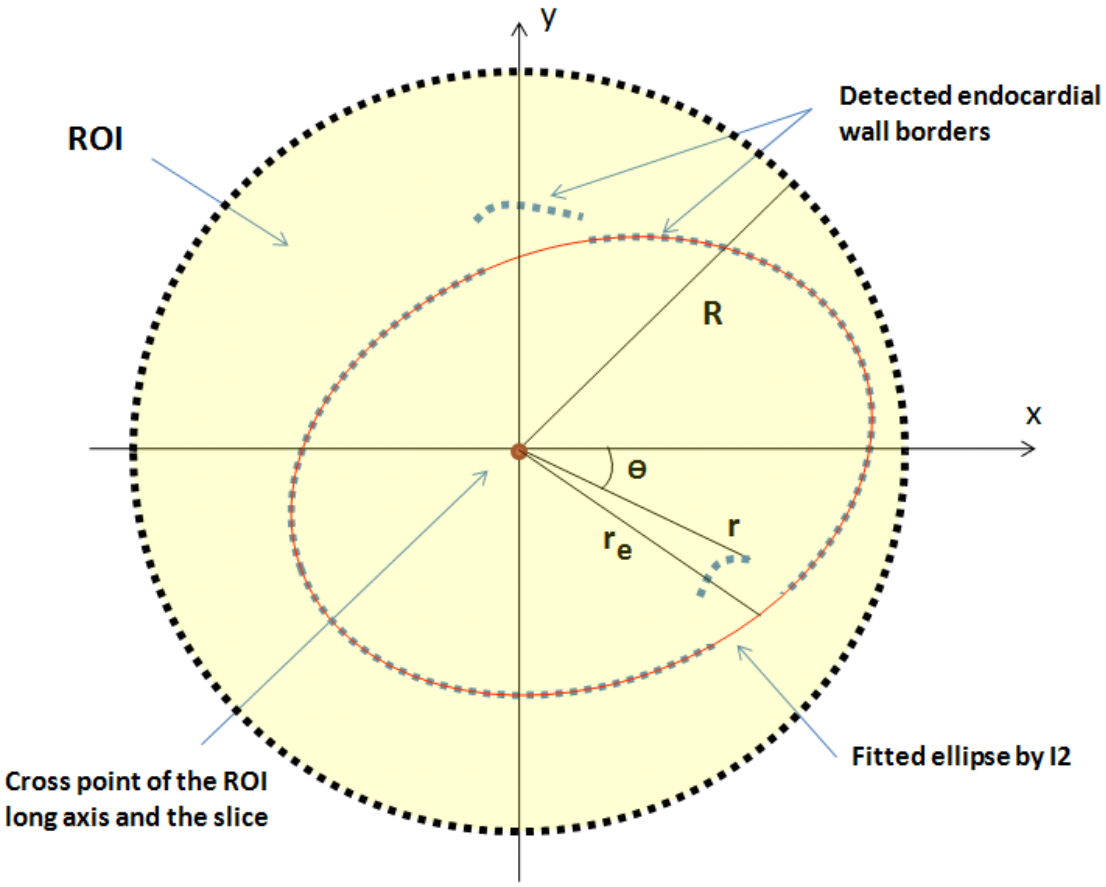


Figure 2-10 A demonstration of the ellipse interpolation algorithm 2 (I_2): A Cartesian coordinate was built up for applying I_2. The blue dots are detected borders. The red ellipse was obtained by the least square method.

After fitting an ellipse into the detected borders, I_2 evaluated the differences between detected radius and fitted ellipse, which was calculated as:

$$D(\theta) = |r(\theta) - r_e(\theta)| \tag{2.13}$$

$$r_e(\theta) = \sqrt{x^2 + y^2} \tag{2.14}$$

where $r(\theta)$ is the detected radius, and $r_e(\theta)$ is the distance from the fitted ellipse to the origin (cross point of the ROI long axis and the slice, Figure 2-10, red point). If $D(\theta)$ was larger than the length of 5 pixels (default setting), the corresponding $r(\theta)$ would be replaced by $r_e(\theta)$

Through border detection and interpolation algorithms, the detected radii $r(\theta)$ were given for all scanned radial lines. However, it is not only the failed detections that need to be interpolated, but also the gap between each two detected neighbouring borders needs to be interpolated, so that the detected cavity inner wall border is a continuous, rather than a discrete contour. This is illustrated in the next section by assuming that each two neighbouring borders together with the cross point construct a small sector.

2.5. Volume calculation

For one 2D short axis slice, r(θ) was obtained by border detection and interpolation algorithms. Then, r(θ) was transferred to pixel indexes so that the LV endocardial wall borders can be indicated in the original slices. If this was repeated for all LV short axis slices in one frame of a 3D echo image, and the detected pixels were registered into 3D Cartesian coordinates, a 3D view of a reconstructed LV endocardial wall surface was obtained frame by frame (Figure 2-11).

To calculate the LV volume, the summation of discs (Simpson's rule) was employed. In Figure 2-11, A, a 2D slice was divided into many small sectors. Each sector was constructed by the radii r(θ), r(θ+Δθ) and slice thickness d. the sector radius was defined as:

$$r_i = [r(\theta) + r(\theta + \Delta\theta)]/2 \tag{2.15}$$

Since the sector volume was:

$$V_i = \frac{1}{2}\Delta\theta \cdot r_i^2 \cdot d \tag{2.16}$$

and the number of sectors was:

$$n = 2\pi/\Delta\theta \tag{2.17}$$

therefore, for one slice, its volume was the summation of all the sectors:

$$V_s = \sum_{i=1}^{i=n} V_i \tag{2.18}$$

Similarly, the LV volume was the summation of all the slice volumes:

$$V_{LV} = \sum_{s=1}^{s=m} V_s \tag{2.19}$$

where m was the number of short axis slices.

Based on this volume calculation algorithm, if one border failed in detection ($r(\theta) = 0$) or detected incorrectly, the volume around that part would also be calculated incorrectly, which resulted in certain bias of the LV volume. Therefore, the volume calculation is very dependent on the quality of the border delineation.

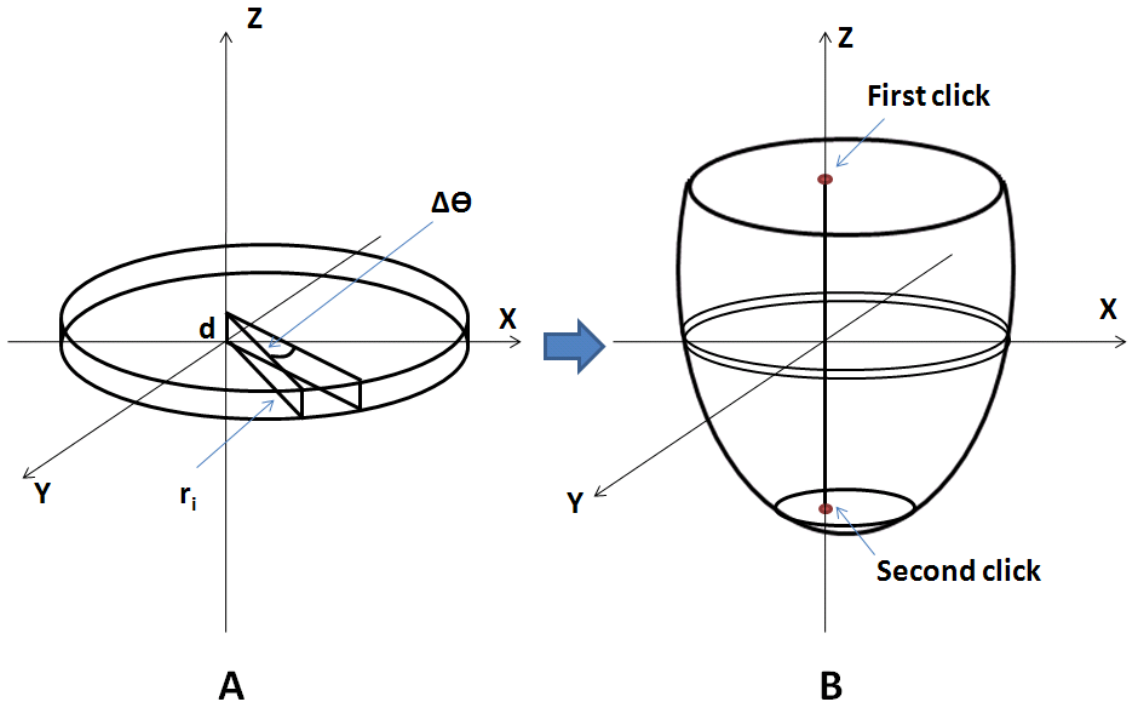


Figure 2-11 3D reconstruction of the LV endocardial wall surface: A. the endocardial wall boundary was obtained in one short axis slice. 'd' is the slice thickness; B. the whole LV surface was constructed slice by slice in a 3D Cartesian coordinates system.

2.6. Finite element model (FEM)

In Section 2.2, the first order derivative gradient operator was introduced. This aimed to alleviate the influence of noise and artefacts in 3D echo images by making use of the 3D operator, so that the image information of a certain voxel was integrated from its surrounding neighbourhood in all directions. After that, in Section 2.4, in order to

improve the automatic border tracing further, the interpolation algorithms (linear or elliptical) were applied in every 2D short axis slice. However, the influence of noise and artefact are inherent in acoustic images, which unavoidably caused incorrect border detection as well as incorrect border interpolation in some of the 2D short axis slices inside the ROI. To overcome this problem, a mathematical model was applied, which was specifically designed for heart ventricle geometry by using 3D finite element meshes referred to a prolate spheroid coordinate system (65).

2.6.1. Mathematical model

The basis of this geometric model is a prolate spheroid, which is a quadric surface in 3D obtained by rotating an ellipse about one of its principal axes. If the ellipse is rotated about its major axis, as shown in Figure 2-12 A, the surface is a prolate spheroid. It gives a good initial approximation to ventricular geometry. After that, this initial 3D surface is deformed and fitted into a traced 3D geometric dataset of a LV endocardial surface.

To simplify the fitting process, the prolate spheroidal coordinate system (λ, θ, μ) was used to describe a prolate spheroid rather than the rectangular Cartesian coordinate system (x, y, z). This is because it permits the use of a linear least-square fitting algorithm in which only the radial λ coordinate is fitted, whereas in the rectangular Cartesian coordinate system, to accomplish the same process, x, y and z coordinates have to be fitted non-linearly.

A prolate spheroid is defined to have a constant radius λ, along with the elevation angle μ (0 to π) and the azimuthal angle θ (0 to 2π), as shown in Figure 2-12 A. A material point (λ, θ, μ) in the endocardial surface described by prolate spheroidal coordinates has the following Cartesian coordinates:

$$x = a\ sinh\lambda\ sin\mu\ cos\theta \tag{2.20}$$

$$y = a\ sinh\lambda\ sin\mu\ sin\theta \tag{2.21}$$

$$z = a\ cosh\lambda cos\mu \tag{2.22}$$

The fixed parameter ‘a’ is the location of the focus on the z-axis (Figure 2-12 A).

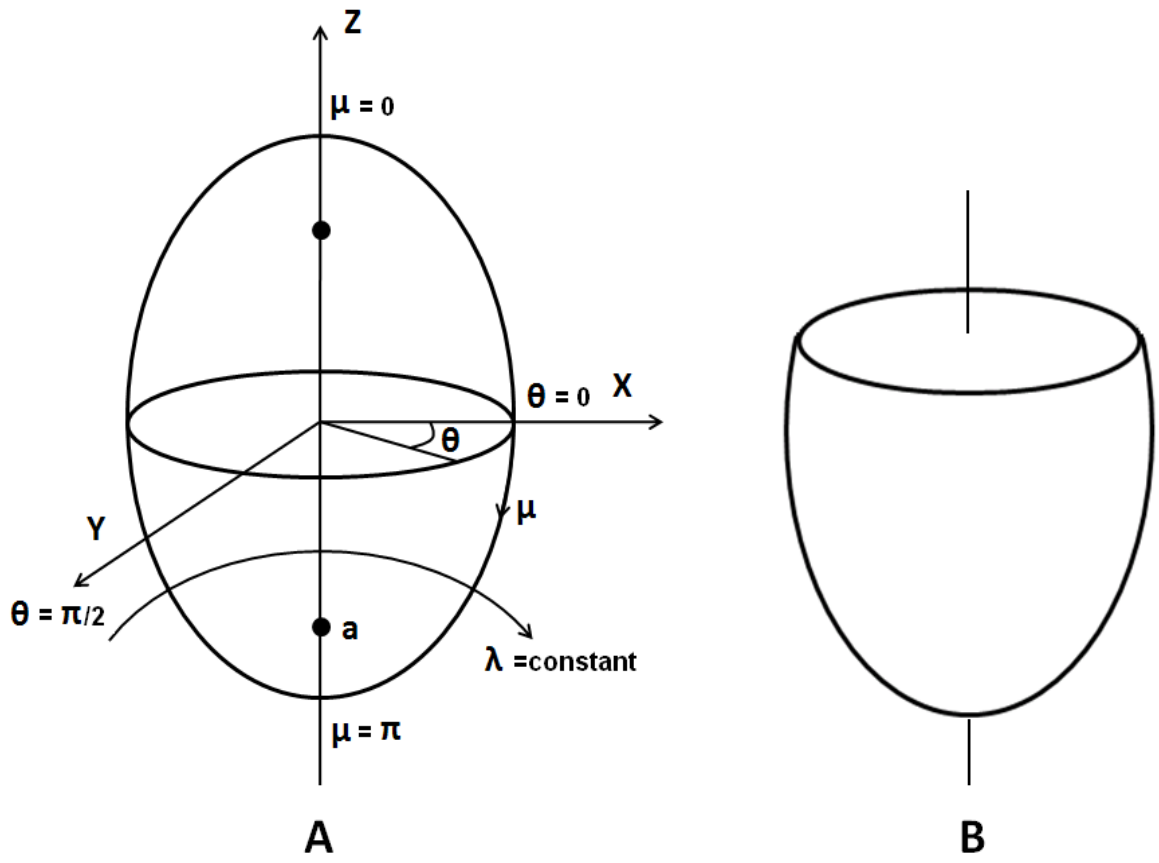


Figure 2-12 A. Prolate spheroid coordinate system (λ, θ, μ) in relation to rectangular coordinate system (x, y, z), and 'a' is the focus point; B. Prolate spheroid coordinate model for the left ventricular endocardial surface (μ∈[π/3, π])

Because the heart left ventricular geometry is not a completed prolate spheroid, the angle μ is normally set within the range of 0 to π. The default range in practice was π/3 to π, as shown in Figure 2-12 B. This model looks like an extremely simplified LV cavity from base to apex.

Table 2-1 Default setting of θ and μ for all the nodes in the finite element mesh

θ	Nodes n	μ	Nodes n
0	1, 5, 9, 13	π/3	1, 2, 3, 4
π	2, 6, 10, 14	5π/9	5, 6, 7, 8
3π/2	3, 7, 11, 15	7π/9	9, 10, 11, 12
2π	4, 8, 12, 16	π	13, 14, 15, 16

Note: node 13, 14, 15 and 16 are overlapped.

After establishing the initial shape of the model, an ensemble of 12 elements with 16 nodes (Figure 2-13 Last 4 nodes are overlapped) was defined to assemble it. The default setting of θ and μ for every node is listed in Table 2-1. Within each element, material coordinates (ξ1, ξ2) lie in circumferential and azimuthal directions, respectively, as shown in Figure 2-13. They both vary from 0 to 1.

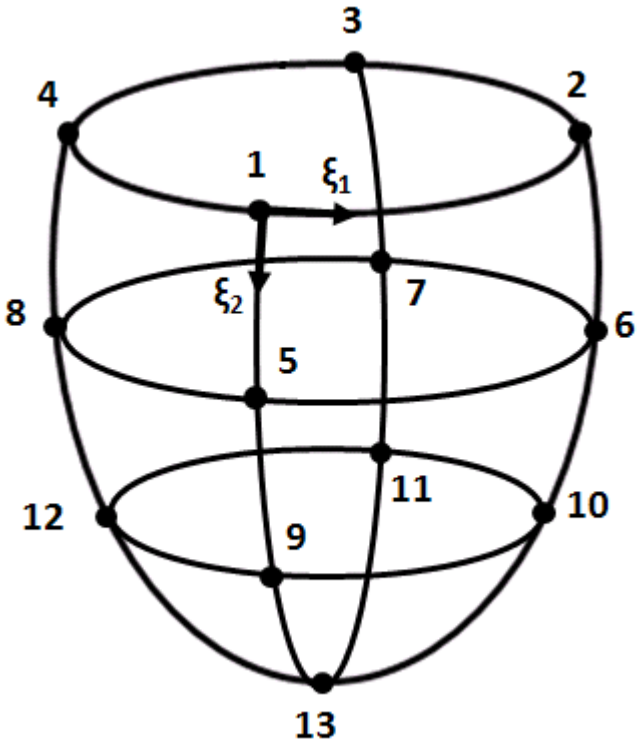


Figure 2-13 Schematic diagram of the finite element mesh of the LV endocardial surface: 12 elements with 16 nodes (number 1 to 13, last 4 nodes are overlapped as the apex); Element material coordinates (ξ1, ξ2) lie in circumferential and azimuthal directions respectively.

The true position of a material point, identified by material coordinates (ξ_1, ξ_2) within an element, is approximated by an interpolation of parameters defined at the element nodes. For linear interpolation, these parameters are simply the values of the coordinates at the nodes (λ_n, θ_n, μ_n), where n = 1, 2, ⋯ 16. In this model, a bilinear interpolation of the nodal values was used for the prolate spheroidal θ-coordinate:

$$\theta(\xi_1,\xi_2) = L_1(\xi_1)L_1(\xi_2)\theta_1 + L_2(\xi_1)L_1(\xi_2)\theta_2 + L_1(\xi_1)L_2(\xi_2)\theta_3 + L_2(\xi_1)L_2(\xi_2)\theta_4 \quad (2.23)$$

where:

$$L_1(\xi) = 1 - \xi \text{ and } L_2(\xi) = \xi$$

They are one-dimensional linear Lagrange basis functions (66).

The μ-coordinate was interpolated in the same way as the θ-coordinate:

$$\mu(\xi_1, \xi_2) = L_1(\xi_1)L_1(\xi_2)\mu_1 + L_2(\xi_1)L_1(\xi_2)\mu_2 + L_1(\xi_1)L_2(\xi_2)\mu_3 + L_2(\xi_1)L_2(\xi_2)\mu_4 \tag{2.24}$$

However, in order to achieve the first-order continuity between elements, i.e. continuity of slope, bilinear interpolation was not capable enough for interpolating the λ-coordinate. So bicubic Hermite basis functions were used in the (ξ_1, ξ_2) surface:

$$\begin{aligned}\lambda(\xi_1, \xi_2) = {} & H_1^0(\xi_1)H_1^0(\xi_2)\lambda_1 + H_2^0(\xi_1)H_1^0(\xi_2)\lambda_2 + H_1^0(\xi_1)H_2^0(\xi_2)\lambda_3 \\ & + H_2^0(\xi_1)H_2^0(\xi_2)\lambda_4 + H_1^1(\xi_1)H_1^0(\xi_2)(\frac{\partial\lambda}{\partial\xi_1})_1 + H_2^1(\xi_1)H_1^0(\xi_2)(\frac{\partial\lambda}{\partial\xi_1})_2 \\ & + H_1^1(\xi_1)H_2^0(\xi_2)(\frac{\partial\lambda}{\partial\xi_1})_3 + H_2^1(\xi_1)H_2^0(\xi_2)(\frac{\partial\lambda}{\partial\xi_1})_4 + H_1^0(\xi_1)H_1^1(\xi_2)(\frac{\partial\lambda}{\partial\xi_2})_1 \\ & + H_2^0(\xi_1)H_1^1(\xi_2)(\frac{\partial\lambda}{\partial\xi_2})_2 + H_1^0(\xi_1)H_2^1(\xi_2)(\frac{\partial\lambda}{\partial\xi_2})_3 + H_2^0(\xi_1)H_2^1(\xi_2)(\frac{\partial\lambda}{\partial\xi_2})_4 \\ & + H_1^1(\xi_1)H_1^1(\xi_2)(\frac{\partial^2\lambda}{\partial\xi_1\partial\xi_2})_1 + H_2^1(\xi_1)H_1^1(\xi_2)(\frac{\partial^2\lambda}{\partial\xi_1\partial\xi_2})_2 \\ & + H_1^1(\xi_1)H_2^1(\xi_2)(\frac{\partial^2\lambda}{\partial\xi_1\partial\xi_2})_3 + H_2^1(\xi_1)H_2^1(\xi_2)(\frac{\partial^2\lambda}{\partial\xi_1\partial\xi_2})_4\end{aligned} \tag{2.25}$$

where:

$$H_1^0(\xi) = 1 - 3\xi^2 + 2\xi^3$$

$$H_1^1(\xi) = \xi(\xi - 1)^2$$

$$H_2^0(\xi) = \xi^2(3 - 2\xi)$$

$$H_2^1(\xi) = \xi^2(\xi - 1)$$

They are one-dimensional cubic Hermite basis functions (66).

The parameters λ, $\frac{\partial\lambda}{\partial\xi_1}$, $\frac{\partial\lambda}{\partial\xi_2}$ and $\frac{\partial^2\lambda}{\partial\xi_1\partial\xi_2}$ defined by bicubic Hermite interpolation functions for each node of the finite element mesh, together with trilinear interpolation functions for parameters θ and μ, which are given above, define the heart ventricular

endocardial surface geometry (65).

2.6.2. Data fitting

By using the semi-automated algorithms described in the previous sections, the LV endocardial surface was traced slice by slice inside the ROI. All traced surface measurement data was reconstructed in a 3D Cartesian coordinate system, as shown in Figure 2 11. Then, they were transferred from the Cartesian coordinate system to the prolate spheroid coordinate system by using Equation 2.20, 2.21 and 2.22. After that, the new established finite element meshes were fitted by using a least squares algorithm.

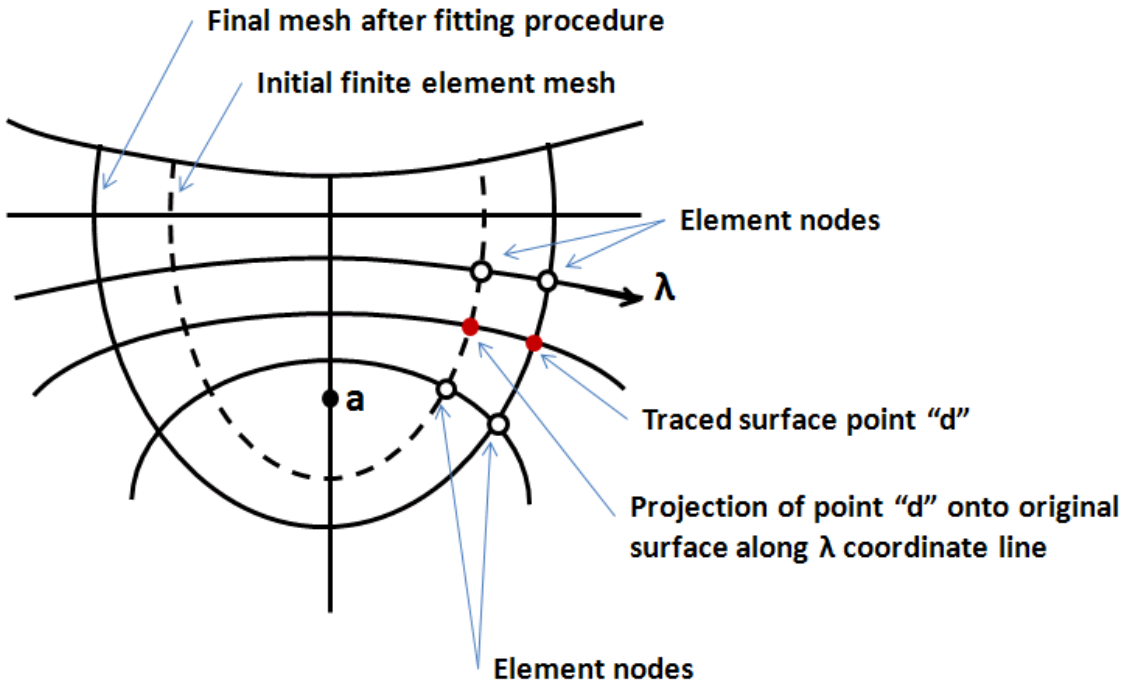


Figure 2-14 Schematic diagram of linear least squares fit of finite element mesh to the traced endocardial surface data set in prolate spheroidal coordinate system; Assumption: a traced LV endocardial surface point "d" and its projection point onto original model surface along λ-coordinate line have the same material coordinates: (ξ_1^d, ξ_2^d). 'a' is the focus point

The choice of a prolate spheroidal coordinate system for describing the geometry of the LV endocardial surface has a major benefit when fitting the mathematical model to the traced surface. The nodal values of θ and μ are held fixed all the time, as given in

Table 2-1. Only the nodal values of λ, $\frac{\partial\lambda}{\partial\xi_1}$, $\frac{\partial\lambda}{\partial\xi_2}$ and $\frac{\partial^2\lambda}{\partial\xi_1\,\partial\xi_2}$ are fitted by using a least squares algorithm, which makes the fitting process much easier than it in the rectangular Cartesian coordinate system.

From the two trilinear interpolation functions of θ-coordinate and μ-coordinate (Equation 2.23 and 2.24), for every traced endocardial surface border, its unique material coordinates (ξ_1^d, ξ_2^d) were calculated in its corresponding element (Calculated two variables in two equations.). Through this way, all the projection points of the traced endocardial surface geometric dataset onto the mathematical model surface along the λ-coordinate were calculated, as shown Figure 2-14.

As introduced before, the major benefit of choosing the prolate spheroidal coordinate system is that only the radial coordinate (λ-coordinate) needs to follow the least squares fitting procedure when fitting the original model surface to the traced endocardial surface dataset. However, there is one assumption of the least squares fitting procedure: the coordinates (ξ_1^d, ξ_2^d) obtained from the projection of the traced surface point d onto the model surface was assumed to be constant during the fitting procedure, although this may not be the truth: The surface coordinates (ξ_1^d, ξ_2^d) changes as the finite element ensemble is moved from the initial configuration to the final shape during minimizing sum of squared projection lengths. However this change has already been minimized by using the prolate spheroidal coordinate system (Figure 2-14).

The linear fitting procedure was obtained by minimizing the sum of squares S:

$$S = \sum_{d=1}^{d=n} \left[\lambda\left(\xi_1^d, \xi_2^d\right) - \lambda_d\right]^2 \tag{2.26}$$

where λ_d is the λ-coordinate of a traced endocardial surface point d; $\lambda\left(\xi_1^d, \xi_2^d\right)$ is the λ-coordinate of the projection of the surface point d onto the model surface along lines of constant θ and μ; n is the number of the traced LV endocardial surface data points (65).

Take one element of the LV endocardial ensemble as an example of minimizing the sum of squares S. The bicubic Hermite interpolation (Equation 2.25) can be written as:

$$\lambda\left(\xi_1^d, \xi_2^d\right) = \Psi_{16}(\xi_1^d, \xi_2^d)\lambda_{16} \tag{2.27}$$

where the matrix:

$$\Psi_{16}\left(\xi_1^{\mathrm{d}}, \xi_2^{\mathrm{d}}\right) = [\mathrm{H}_1^0\left(\xi_1^{\mathrm{d}}\right)\mathrm{H}_1^0\left(\xi_2^{\mathrm{d}}\right) \quad \mathrm{H}_2^0\left(\xi_1^{\mathrm{d}}\right)\mathrm{H}_1^0\left(\xi_2^{\mathrm{d}}\right) \quad \cdots \quad \mathrm{H}_2^1\left(\xi_1^{\mathrm{d}}\right)\mathrm{H}_2^1\left(\xi_2^{\mathrm{d}}\right)] \tag{2.28}$$

is formed by the 16 basis functions of the bicubic Hermite interpolation function (Equation 2.25). The matrix:

$$\lambda_{16} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \left(\dfrac{\partial^2 \lambda}{\partial \xi_1^d \, \partial \xi_2^d}\right)_3 \\ \left(\dfrac{\partial^2 \lambda}{\partial \xi_1^d \, \partial \xi_2^d}\right)_4 \end{bmatrix} \tag{2.29}$$

is formed by the associated 16 nodal values. Therefore, the Equation 2.26 becomes:

$$S = \sum_{\mathrm{d}=1}^{d=n} \left[\Psi_{16}\left(\xi_1^d, \xi_2^d\right)\lambda_{16} - \lambda_d\right]^2 \tag{2.30}$$

To minimize S, $S' = 0$

$$\therefore 2\sum_{1}^{n} \Psi_{16}^T[\Psi_{16}(\xi_1^d, \xi_2^d)\lambda_{16} - \lambda_d] = 0 \tag{2.31}$$

where Ψ_{16}^T is the transpose of the matrix $\Psi_{16}\left(\xi_1^d, \xi_2^d\right)$.

$$\therefore \lambda_{16} = (A)^{-1}B \tag{2.32}$$

where

$$A = \sum_{1}^{n} \Psi_{16}^T \Psi_{16} \tag{2.33}$$

$$B = \sum_{1}^{n} \Psi_{16}^T \lambda_d \tag{2.34}$$

For every node of the finite element mesh, there are four parameters of λ-coordinate

given by the bicubic Hermite interpolation: ($\lambda_n, \frac{\partial \lambda_n}{\partial \xi_1}, \frac{\partial \lambda_n}{\partial \xi_2}$ and $\frac{\partial^2 \lambda_n}{\partial \xi_1 \partial \xi_2}$), where n = 1, 2,⋯16. Therefore in total there are 64 parameters for 16 nodes (where the last 4 nodes are overlapped in order to construct the LV apex in this case, so the values of the last 16 parameters are the same). During the least squares fitting algorithm, all 64 nodal parameters were obtained for a certain measured data set by minimizing the sum of squares S (Equation 2.26 to 2.34), which meant the 3D mathematical model was adapted to the semi-automated traced LV endocardial surface. Finally, the volume of the 3D finite element mesh was calculated by the same way described in Section 2.5. The default setting of the material coordinates (ξ_1, ξ_2) is increasing from 0 to 1 each time 0.01 step by step. All the material points were transferred to the rectangular Cartesian coordinates system (Equation 2.20, 2.21 and 2.22), and the volume was calculated by the summation of discs method (Simpson's rule, Equation 2.19) finally.

2.7. Summary

In this chapter, the newly developed semi-automated algorithms were introduced. They traced the LV endocardial wall surface and calculated the LV volume quantitatively. The mathematical finite element model was also introduced, which fitted a 3D prolate spheroidal surface into a semi-automated traced LV endocardial surface, and calculated the LV volume.

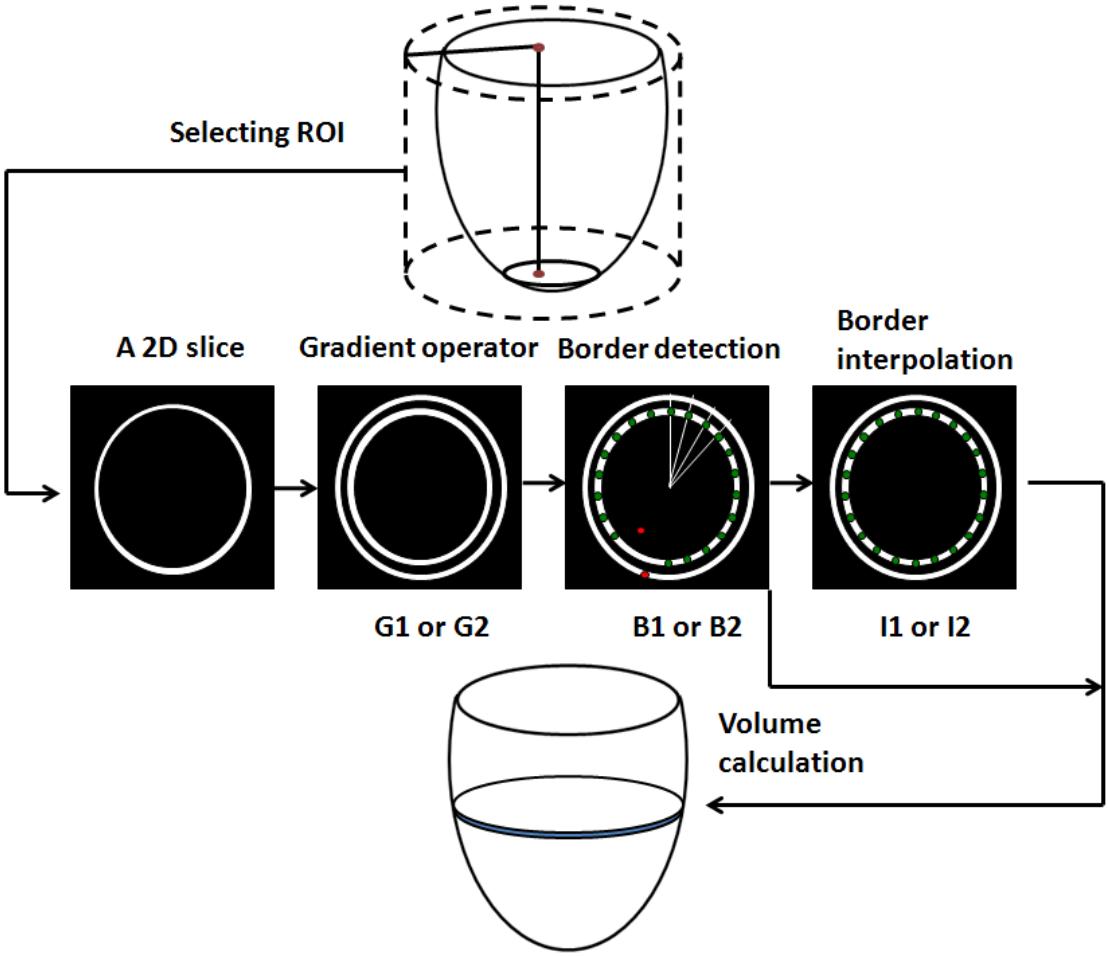


Figure 2-15 Summary of the semi-automated algorithms

As shown in Figure 2-15, the basic procedure of semi-automated algorithms was: ROI selection, applying gradient operator, border detection, border interpolation (with or without) and volume calculation. To better describe the algorithm, some steps of the procedure were coded by the combination of a letter and a number as follows:

G_1: 2D gradient operator;
G_2: 3D gradient operator;
B_1: border detection algorithm 1;

B_2: border detection algorithm 2;
I_1: linear interpolation algorithm;
I_2: elliptical interpolation algorithm;

By combining them in different orders, 12 different methods for tracing borders and calculating the LV volume were obtained, which are listed in Table 2-2.

Table 2-2 12 algorithms generated from the combination of edge operators, border detection and interpolation algorithms

Algorithms	Gradient operator	Border detection	Border Interpolation
	G_1 = 2D	**B_1 = detection 1**	**I_1 = linear**
	G_2 = 3D	**B_2 = detection 2**	**I_2 = elliptical**
G_1B_1	G_1	B_1	—
G_1B_2	G_1	B_2	—
$G_1B_1I_1$	G_1	B_1	I_1
$G_1B_2I_1$	G_1	B_2	I_1
$G_1B_1I_2$	G_1	B_1	I_2
$G_1B_2I_2$	G_1	B_2	I_2
G_2B_1	G_2	B_1	—
G_2B_2	G_2	B_2	—
$G_2B_1I_1$	G_2	B_1	I_1
$G_2B_2I_1$	G_2	B_2	I_1
$G_2B_1I_2$	G_2	B_1	I_2
$G_2B_2I_2$	G_2	B_2	I_2

To assess the performance of the 12 semi-automated methods and the finite element model for heart left ventricular endocardial surface delineation and volume quantification, a series of analyses and tests were designed and assessed. Starting from the simple laboratory balloon phantom (Chapter 3), to the more complicated and accurate tissue-mimicking phantom (Chapter 4), the accuracy and reproducibility of different algorithms were evaluated and compared. After that, the FEM was applied to the traced geometric dataset from the best algorithm, and its performance compared. Meanwhile, border detection and interpolation algorithms were validated by 2D magnetic resonance images (MRI, Chapter 5). Finally, all the algorithms were applied to the clinical 3D echo images of human subjects (Chapter 6).

Chapter 3. Quantitative volume measurement of balloon phantoms

Chapter 2 described the development of new semi-automated algorithms for LV endocardial surface delineation and LV volume calculation. We also introduced the mathematical finite element model (FEM) that can make full use of traced geometric data to reconstruct the LV endocardial surface in 3D. In this chapter, the performance of all algorithms is assessed by applying them to 3D echo images of laboratory phantoms.

The most important two issues of human heart chamber volume quantification are accuracy and reproducibility. However, in the clinical environment, it is very difficult to evaluate the accuracy of different volume quantification techniques, because there is no widely accepted gold standard in this field. For non-invasive measurement of the left ventricular volume of the human heart, manual tracing of a stack of short axis images plus Simpson's rule (disc summation method) are normally considered as the standard, but there are also some drawbacks affecting the accuracy in this technique. For example, the manually traced borders show significant inter- and intra-operator variance. The short axis images are discrete slices of fixed thickness, so the LV endocardial surface is not fully traced. Therefore, to assess the accuracy and the reproducibility of our newly developed techniques, laboratory phantoms with a controlled volume imaged by 3D echocardiography were adopted.

As discussed in Chapter 1, comparing with conventional 2D echocardiography, it has been widely believed that 3D echocardiography can provide better performance for quantitative volume measurement. However, the conventional analysis of the 3D echo images still utilizes information of 3D images by multiple separate 2D slices. In other words, there is no combination of the 3D image information from all three dimensions while analyzing a 3D image. Therefore, one hypothesis is that, if 3D echo images can be analyzed in a 3D way, a better result should be achieved. In this case, the 3D gradient operator and the 3D FEM model should provide better accuracy as well as better reproducibility in quantitative volume measurements of 3D echo images.

3.1. Imaging acquisition

A balloon phantom filled with water was used as the imaging subject. The balloon modelled the LV myocardium with water inside mimicking the blood. The phantom was suspended in a water tank, and a Philips Sonos 7500 3D ultrasound system was used to scan it from above. Nine fill volumes (25ml, 50ml, 75ml, 100ml, 150ml, 200ml, 300ml, 400ml and 500ml) were scanned twice sequentially, which resulted in two 3D movies (dynamic 3D images showing static subjects) for each balloon volume. Each 3D movie consisted of a sequence of 3D images (18-20 images/phases). Then all the movies were transferred to a computer for off-line image analysis.

Acknowledgement: we thank Joanne Wild and James Pemberton who collected these images in their previous project.

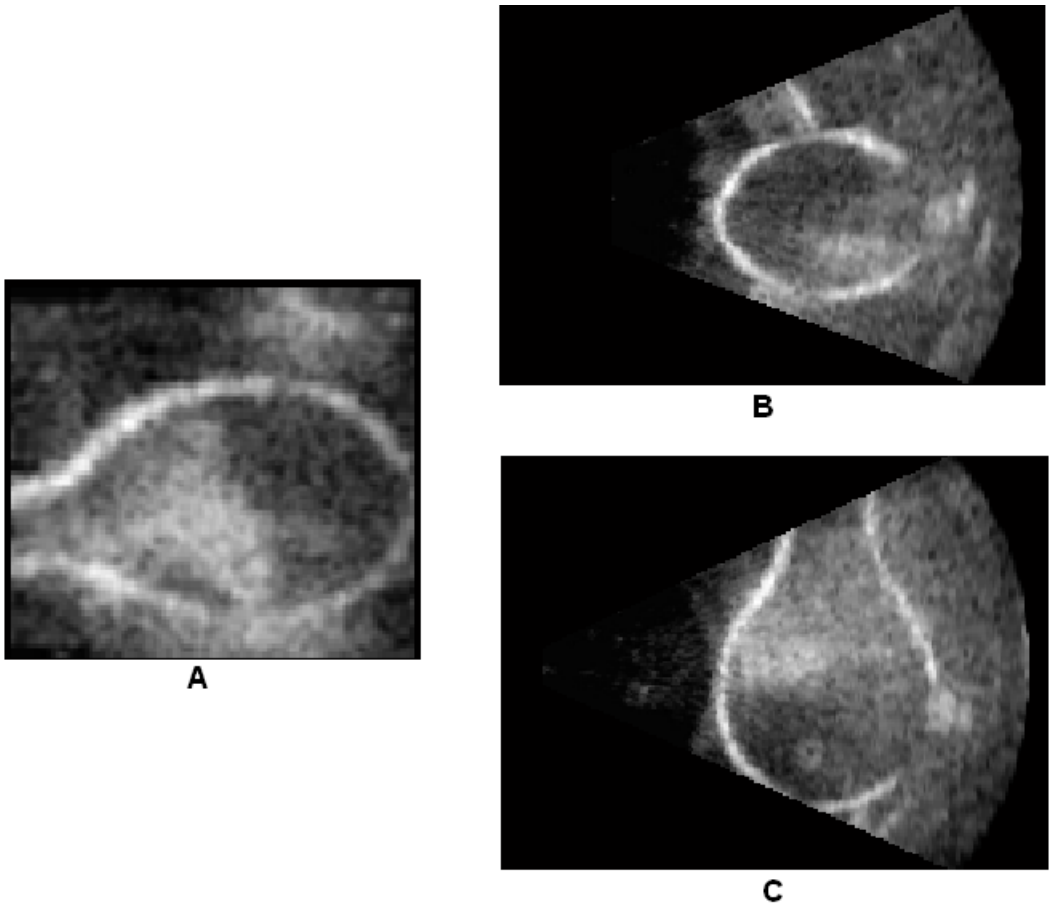


Figure 3-1 2D slices extracted from one 3D echo image of 75ml balloon phantom: A. x-y plane; B. y-z plane; C. z-x plane

3.2. Image analysis

The 3D movies were saved in DICOM (Digital Imaging and Communications in Medicine) format, and all 3D images (all phases) had the same number of voxels (x×y×z: 144×160×208). The semi-automated algorithms were applied on the first five 3D images from one of the 3D movies of each balloon size to measure the volume. The software used for developing the algorithms and analysing the images was Matlab 2008a (The MathWorks, Inc.). The ROI selection Method 1 (Chapter 2, section 2.1) was used to establish the region of interest for all the measurements. In Method 1, two points were selected manually on two extracted short axis slices (one apical and one basal slice) respectively, to establish one ROI. All manual initialization was performed by one individual.

Figure 3-1 shows an example of the gray scale image quality of the 3D echocardiography. These three pictures were extracted from one 3D echo image of the 75ml balloon. The white contour is the balloon wall. Similar to images acquired from conventional 2D echocardiography, because of noise and artefact, the spatial resolution of the 3D echo images was sometimes poor in all three pictures. Some parts of the balloon wall were missing, and some areas inside the balloon, which are normally dark, were almost as bright as the balloon wall contour.

3.2.1. Operator selection

To alleviate the influence of the noise in volume measurement, first order derivative gradient operators were applied to the 3D echo images, so that the original intensity images were transferred into gradient images.

Two types of edge operators were introduced in Chapter 2.2: 2D and 3D gradient operator (G_1 and G_2). Because different sizes of edge operators provide different edge enhancement and noise reduction effects, two G_1 operators (2D), 11×11 and 5×5, were chosen to compare with one 5×5×5 G_2 operator (3D). The first G_1 had almost the same number of pixels/voxels as the G_2 (121 pixels vs. 125 voxels). The second G_1 has the same number of pixels/voxels as one plane of the G_2 (25 vs. 25 pixels).

After applying the edge operators, the border detection algorithms (B_1 and B_2) and interpolation algorithms (I_1 and I_2) were applied. The phantom inner wall surface was then reconstructed, and the balloon volume was calculated. The performance of all combinations of different algorithms was then determined. After that the finite element model was fitted into the geometric data traced by the algorithm with the best performance to calculate the phantom volume.

3.2.2. Ultrasound speed correction

The ultrasound propagation speed depends on mass of molecules, spacing of molecules and force between molecules. Therefore, it varies in different materials. The speed of ultrasound in water (1433m/s) is different from human soft tissue and blood (average 1540m/s), which is also the default setting in 2D and 3D echocardiography. Since the balloon phantom was filled with water, and suspended in a water tank while acquiring the images, there was an error between measured distance and actual distance caused by the difference of speed. This error was calculated as below:

$$V_{water} = 1433m/s;\ V_{tissue} = 1540m/s$$

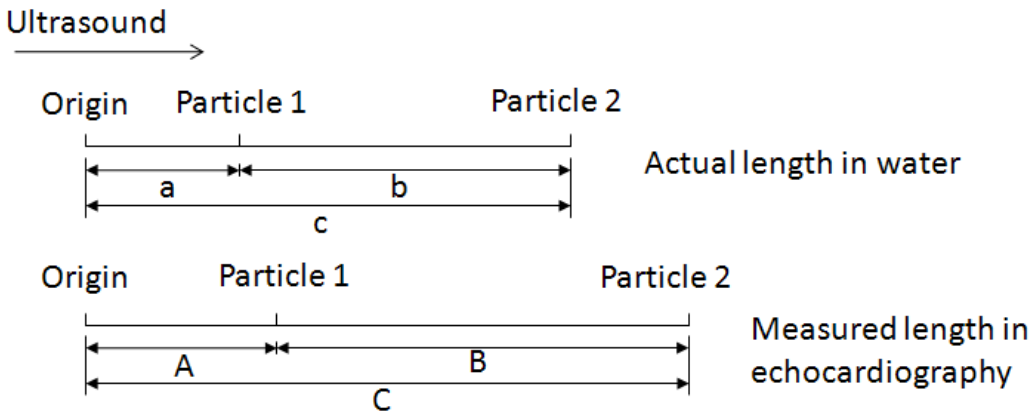


Figure 3-2 Difference between actual distance b in water and measured distance B in echocardiography

As shown in Figure 3-2, ultrasound travelled in water from the origin to particle 1 and particle 2. If Ta and Tc represented the time of ultrasound travelling in distance a and c respectively, then

$$a = V_{water} \cdot Ta \text{ and } c = V_{water} \cdot Tc$$

$$b = c - a = V_{water} \cdot (Tc - Ta)$$

However, the measured distance would be:

$$A = V_{tissue} \cdot Ta \; and \; c = V_{tissue} \cdot Tc$$

$$B = C - A = V_{tissue} \cdot (Tc - Ta)$$

Therefore,

$$\frac{b}{B} = \frac{V_{water}}{V_{tissue}} = \frac{1433m/s}{1540m/s} = 0.93$$

The actual distance b = 0.93B.

This was reflected to the voxel dimensions. For every voxel of a 3D echo image, the actual length in x, y and z dimension was the measured length times 0.93.

3.2.3. Statistical analysis

To distinguish two 2D operators, the 5×5 and 11×11 operators are coded as $G_{1.1}$ and $G_{1.2}$. The 5×5×5 3D operator is still coded as G_2. After the combination with border detection and interpolation algorithms, there were 18 algorithms in total to measure the balloon phantom volumes, as shown in Table 3-1.

Table 3-1 18 algorithms generated from the combination of edge operators, border detection and interpolation algorithms

		Border detections					
		B_1	**B_2**	**B_1**	**B_2**	**B_1**	**B_2**
Interpolations		**None**		**I_1**		**I_2**	
Edge operators	**$G_{1.1}$**	$G_{1.1}B_1$	$G_{1.1}B_2$	$G_{1.1}B_1I_1$	$G_{1.1}B_2I_1$	$G_{1.1}B_1I_2$	$G_{1.1}B_2I_2$
	$G_{1.2}$	$G_{1.2}B_1$	$G_{1.2}B_2$	$G_{1.2}B_1I_1$	$G_{1.2}B_2I_1$	$G_{1.2}B_1I_2$	$G_{1.2}B_2I_2$
	G_2	G_2B_1	G_2B_2	$G_2B_1I_1$	$G_2B_2I_1$	$G_2B_1I_2$	$G_2B_2I_2$

Note: G1.1 is the 5×5 edge operator, and G1.2 is the 11×11 edge operator.

Since there were nine different volumes of the balloon phantoms, and for each volume, five 3D echo images (five frames of a movie) were selected to apply the 18 semi-automated algorithms, overall, there were 9×5×18 = 810 volume measurements. The reason for selecting five images of each balloon volume was to access the reproducibility of the algorithms. Because the balloon phantom stayed still during the imaging acquisition process, the five 3D echo images were the same. However, the

manual initialization was performed separately for each image so that the ROIs of five images were established differently. Therefore, the variation of the five measurements indicated algorithm reproducibility was achieved.

In order to evaluate the accuracy of all 18 algorithms, the Bland Altman (67) method was applied by comparing the volumes calculated by each algorithm with the known balloon volumes. The agreement was expressed by bias (mean difference), 95% confidence interval and the limits of agreement (2×SD around the mean difference). Besides the Bland Altman method, the ANOVA (analysis of variance) was also employed to compare each category of the algorithms ($G_{1.1}$ vs. $G_{1.2}$ vs. G_2, B_1 vs. B_2, no interpolation vs. I_1 vs. I_2) in accuracy.

In order to evaluate the reproducibility of the algorithms, the coefficient of variation (CV) of each group of five volumes from the five selected repeat 3D echo images measured by each algorithm was calculated (810/5 = 162 calculation of CV). After that, the ANOVA was applied to compare each category of the algorithms ($G_{1.1}$ vs. $G_{1.2}$ vs. G_2, B_1 vs. B_2, no interpolation vs. I_1 vs. I_2) in reproducibility.

By comparing the accuracy and reproducibility, the best algorithm was chosen, and the FEM was applied to the geometric data given by it. Then the Bland Altman and student t test were used to access whether the FEM was able to further improve the accuracy and reproducibility of the volume measurement or not.

3.3. Results

The 18 semi-automated algorithms all succeeded in measuring the balloon volumes once the ROIs were established. It took less than 10 seconds to establish a ROI, trace a balloon inner wall surface and calculate a volume. The results are presented below:

3.3.1. Results of accuracy assessment

After the Bland Altman comparisons, the performance of each semi-automated algorithm in accuracy was shown in Table 3-2. Figure 3-3 and Figure 3-4 shows Bland Altman plots comparing known volumes (VK) of balloon phantom with measured

volumes given by the 18 algorithms for absolute (ml) and relative ((Measured volume - Phantom volume)/Phantom volume, %) values. Table 3-2 shows that the volumes measured by all semi-automated algorithms were significantly different from the known balloon phantom volumes. This is also shown in Figure 3-3 and Figure 3-4, in which all the 95% CI (dash lines) of the biases did not cross the zero axes. The biases were all positive, which meant that semi-automated algorithms underestimated balloon volumes. The biases in relative values indicated that the balloon volumes were underestimated by approximately 19% to 39%.

Table 3-2 Summary of Bland Altman comparisons (bias±SD) between known balloon phantom volumes and measured volumes from 18 semi-automated algorithms, in both absolute (ml) and relative values (%), respectively

	Measured volume − Phantom volume (ml)					
	B_1	**B_2**	**B_1I_1**	**B_2I_1**	**B_1I_2**	**B_2I_2**
$G_{1.1}$	-55.5±44.6*	-62.2±46.4*	-32.8±26.5*	-37.4±26.7*	-29.7±22.0*	-36.9±25.2*
$G_{1.2}$	-61.2±48.1*	-70.6±54.1*	-37.5±28.2*	-46.2±33.7*	-34.7±24.3*	-45.6±30.8*
G_2	-53.7±41.7*	-60.5±42.9*	-30.4±21.5*	-37.1±24.2*	-27.0±17.7*	-36.0±21.8*
	(Measured volume − Phantom volume)/Phantom volume (%)					
	B_1	**B_2**	**B_1I_1**	**B_2I_1**	**B_1I_2**	**B_2I_2**
$G_{1.1}$	-31.6±8.6*	-35.3±7.9*	-19.6±6.5*	-22.0±6.1*	-19.5±8.3*	-22.8±8.1*
$G_{1.2}$	-35.1±9.3*	-39.2±8.2*	-22.0±6.4*	-25.8±6.7*	-22.0±8.0*	-26.8±7.0*
G_2	-31.2±8.9*	-34.9±8.0*	-19.0±6.7*	-22.2±6.3*	-18.6±8.2*	-22.8±7.8*

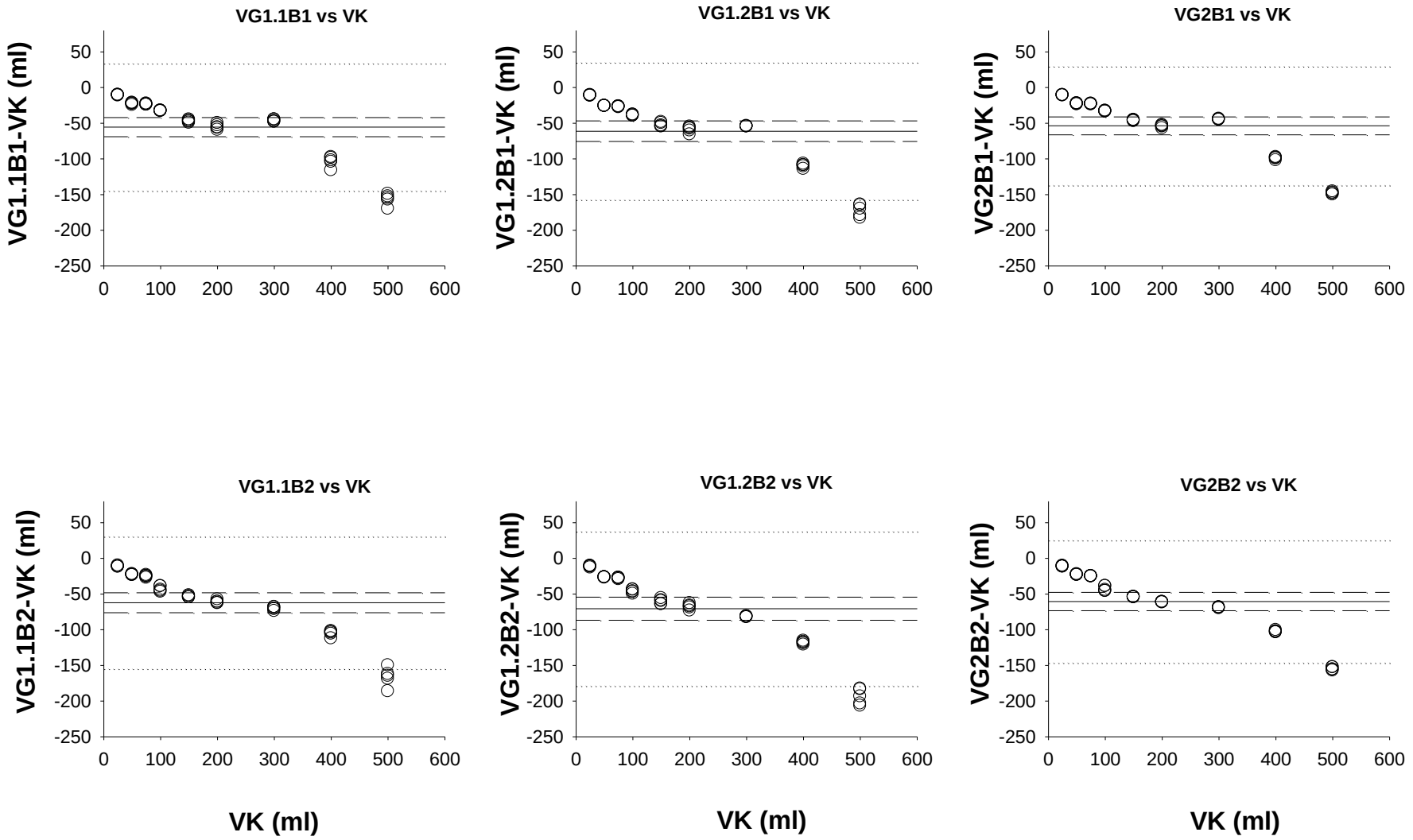
Note: '*' means significant difference.

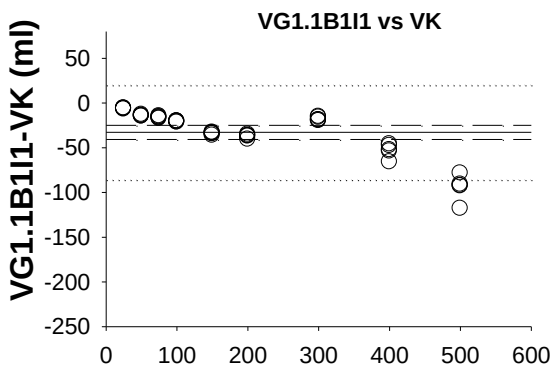
The ANOVA was applied to the differences between known balloon volumes and measured volumes (Vmeasured - Vknown) in order to analyze the performance of each step of the semi-automated algorithms (gradient operator, border detection and interpolation) in accuracy.

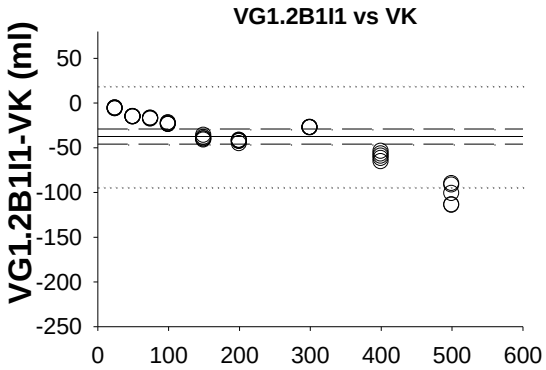
For three edge operators, the results showed that there was no significant difference between $G_{1.1}$ and G_2 (P=0.57; mean bias with Vknown: $G_{1.1}$: 42.4ml and G_2: 40.8ml). However, there was a significant difference between $G_{1.1}$ and $G_{1.2}$ (P<0.05), as well as G_2 and $G_{1.2}$ (P<0.05; mean bias with Vknown: $G_{1.2}$: 49.4ml). For two border detection algorithms, the results showed significant difference (P<0.05; mean bias with Vknown: B_1: 40.2ml and B_2: 48.1ml). For interpolation algorithms, although no significant difference was found between I_1 and I_2 (P=0.51; mean bias with Vknown: I_1: 36.9ml

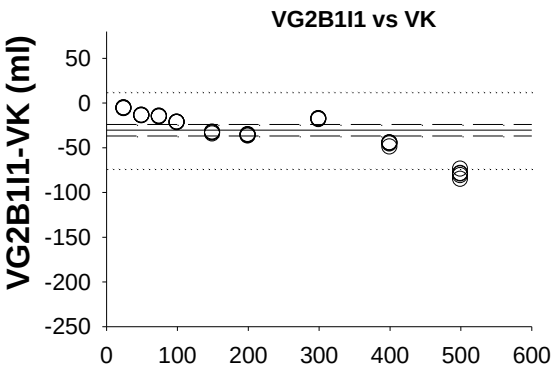
and I_2: 34.9ml), there was a significant difference between with and without interpolations ($P<0.05$; mean bias with Vknown: no interpolation: 60.6ml).

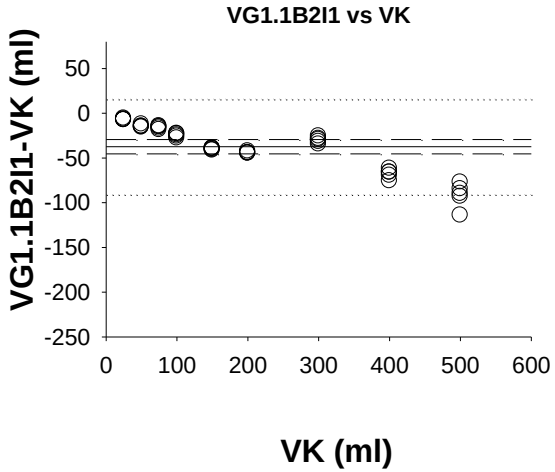
Overall, for accuracy assessment, the edge operator $G_{1.1}$ and G_2 showed better results than $G_{1.2}$. The border detection algorithm B_1 showed a significantly smaller bias than B_2. After applying the interpolation algorithms (I1 or I2), the results showed similar biases for two different interpolations. The interpolation algorithms made a great improvement to the border detection algorithms. The absolute mean bias was reduced from 61ml (35%) to less than 37ml (21%).

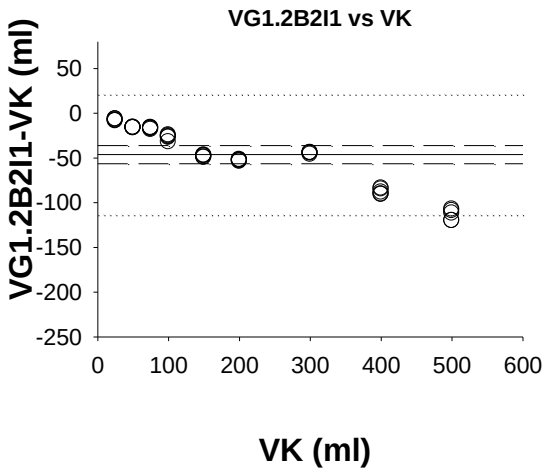

VG1.1B1 vs VK
VG1.2B1 vs VK
VG2B1 vs VK
VG1.1B2 vs VK
VG1.2B2 vs VK
VG2B2 vs VK
VG1.1B1-VK (ml)
VG1.2B1-VK (ml)
VG2B1-VK (ml)
VG1.1B2-VK (ml)
VG1.2B2-VK (ml)
VG2B2-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600
VK (ml)

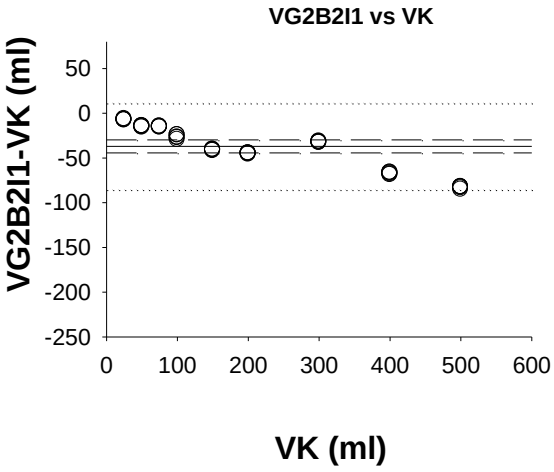

VG1.1B1I1 vs VK
VG1.1B1I1-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600


VG1.2B1I1 vs VK
VG1.2B1I1-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600


VG2B1I1 vs VK
VG2B1I1-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600


VG1.1B2I1 vs VK
VG1.1B2I1-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600
VK (ml)


VG1.2B2I1 vs VK
VG1.2B2I1-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600
VK (ml)


VG2B2I1 vs VK
VG2B2I1-VK (ml)
50
0
-50
-100
-150
-200
-250
0
100
200
300
400
500
600
VK (ml)

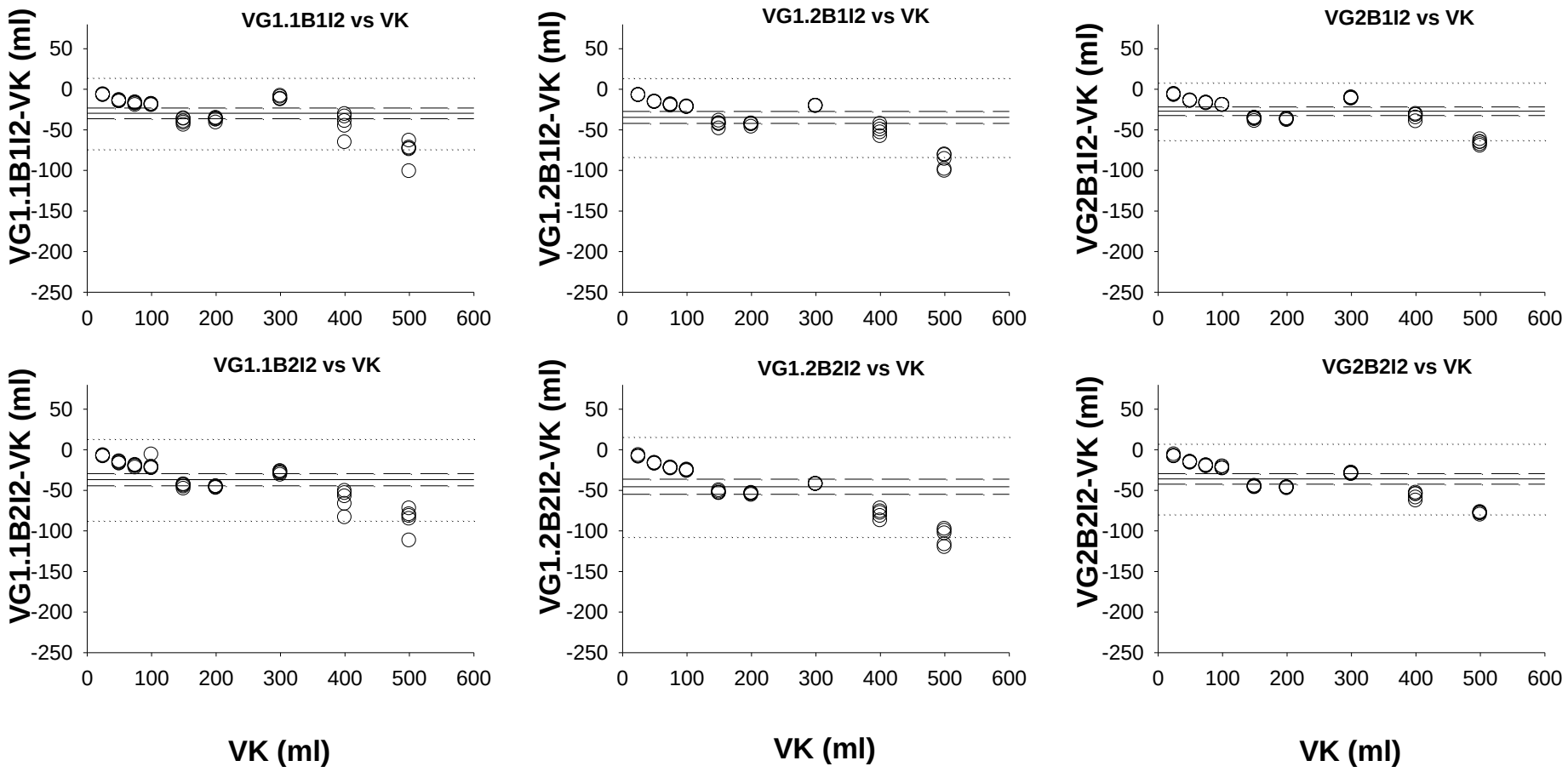
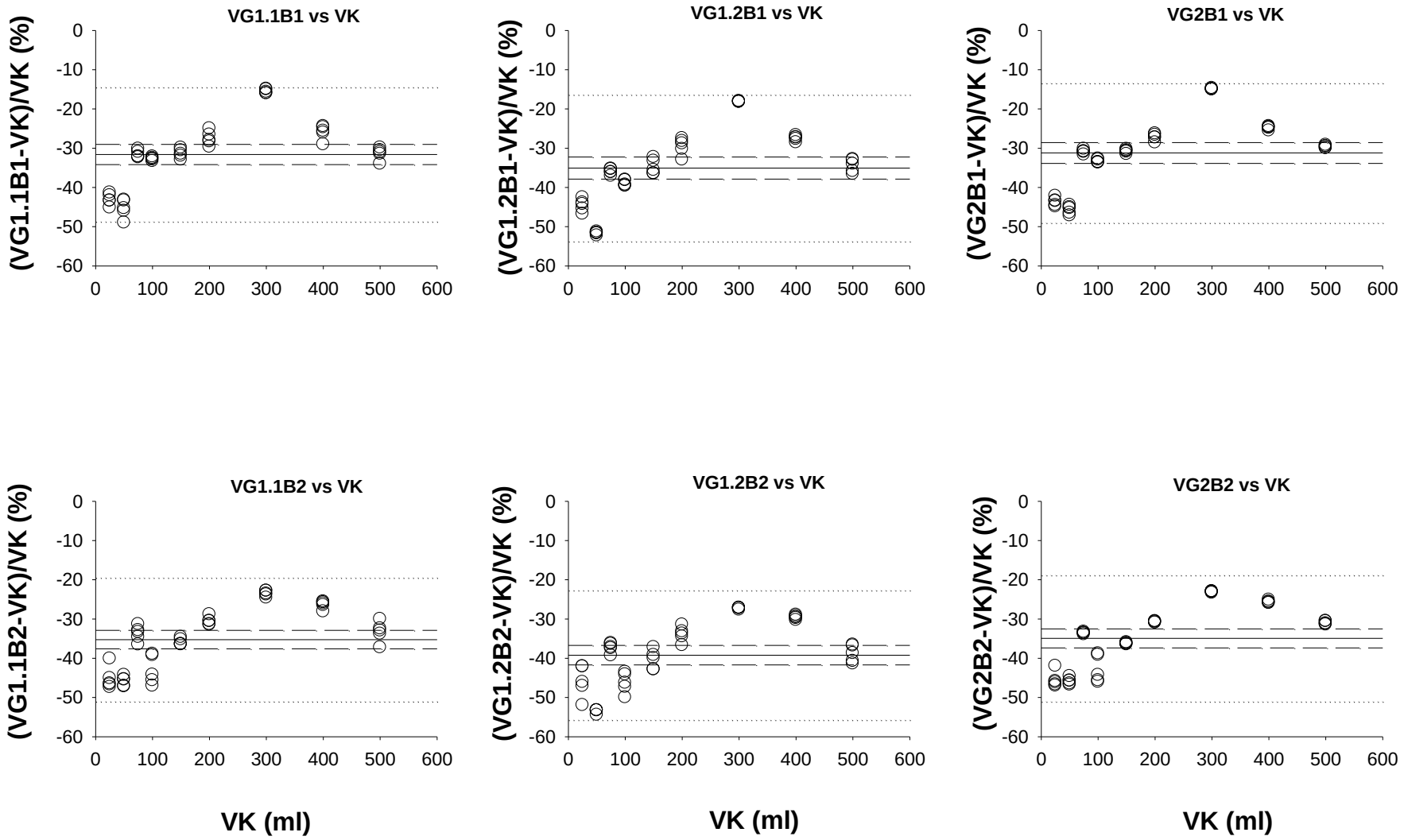
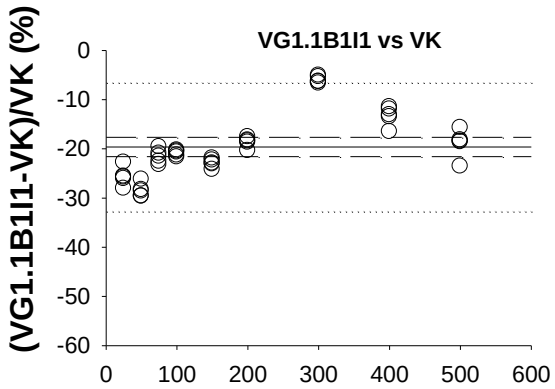
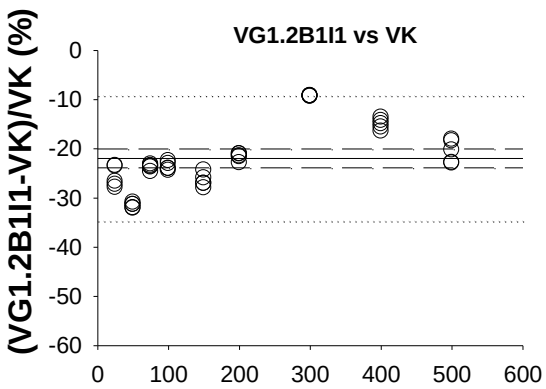
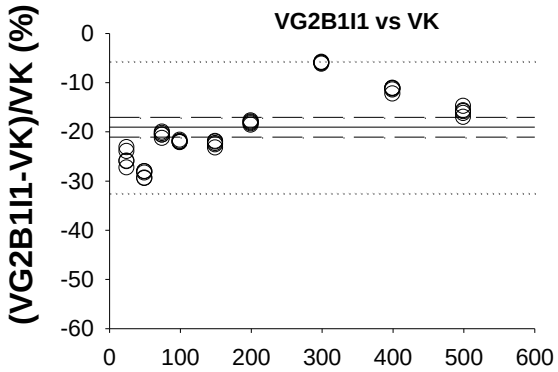


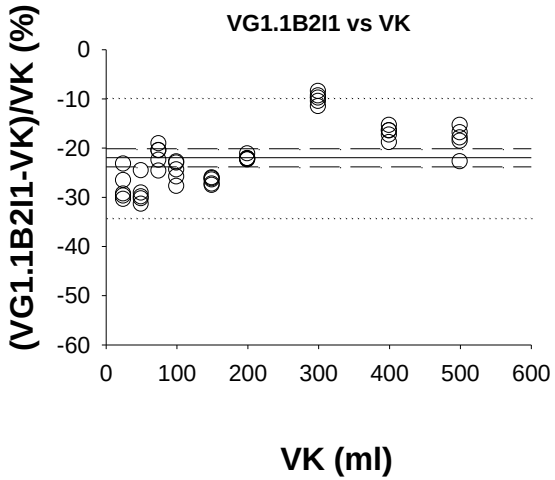
Figure 3-3 Bland Altman plots for 18 algorithms: Comparison between Known Volumes (VK) of balloon phantom and Measured Volumes (ml). Solid lines: bias (mean difference); dash lines: 95% confidence interval (CI) of bias; dotted lines: upper and lower limits of agreement

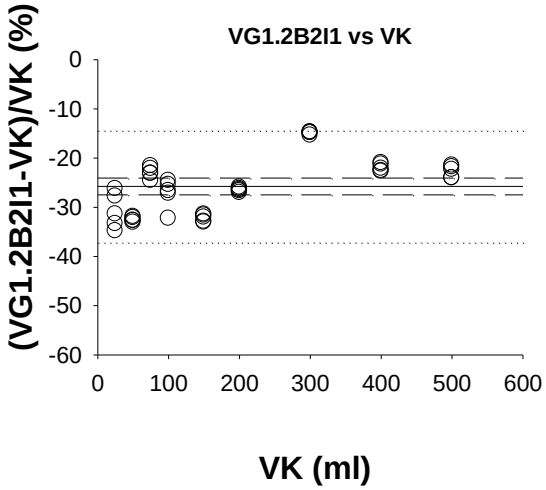

VG1.1B1 vs VK
VG1.2B1 vs VK
VG2B1 vs VK
VG1.1B2 vs VK
VG1.2B2 vs VK
VG2B2 vs VK
(VG1.1B1-VK)/VK (%)
(VG1.2B1-VK)/VK (%)
(VG2B1-VK)/VK (%)
(VG1.1B2-VK)/VK (%)
(VG1.2B2-VK)/VK (%)
(VG2B2-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600
VK (ml)

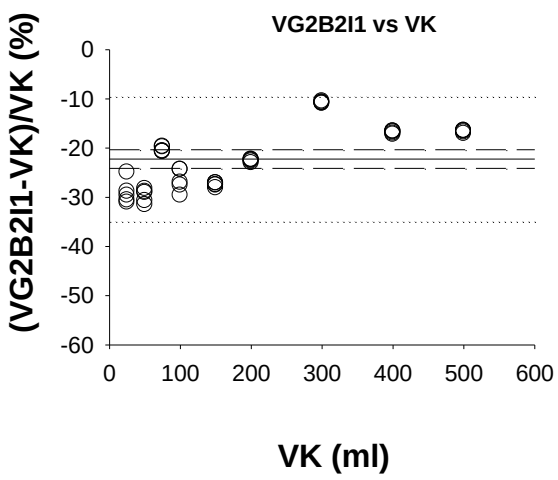

VG1.1B1I1 vs VK
(VG1.1B1I1-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600


VG1.2B1I1 vs VK
(VG1.2B1I1-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600


VG2B1I1 vs VK
(VG2B1I1-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600


VG1.1B2I1 vs VK
(VG1.1B2I1-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600
VK (ml)


VG1.2B2I1 vs VK
(VG1.2B2I1-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600
VK (ml)


VG2B2I1 vs VK
(VG2B2I1-VK)/VK (%)
0
-10
-20
-30
-40
-50
-60
0
100
200
300
400
500
600
VK (ml)

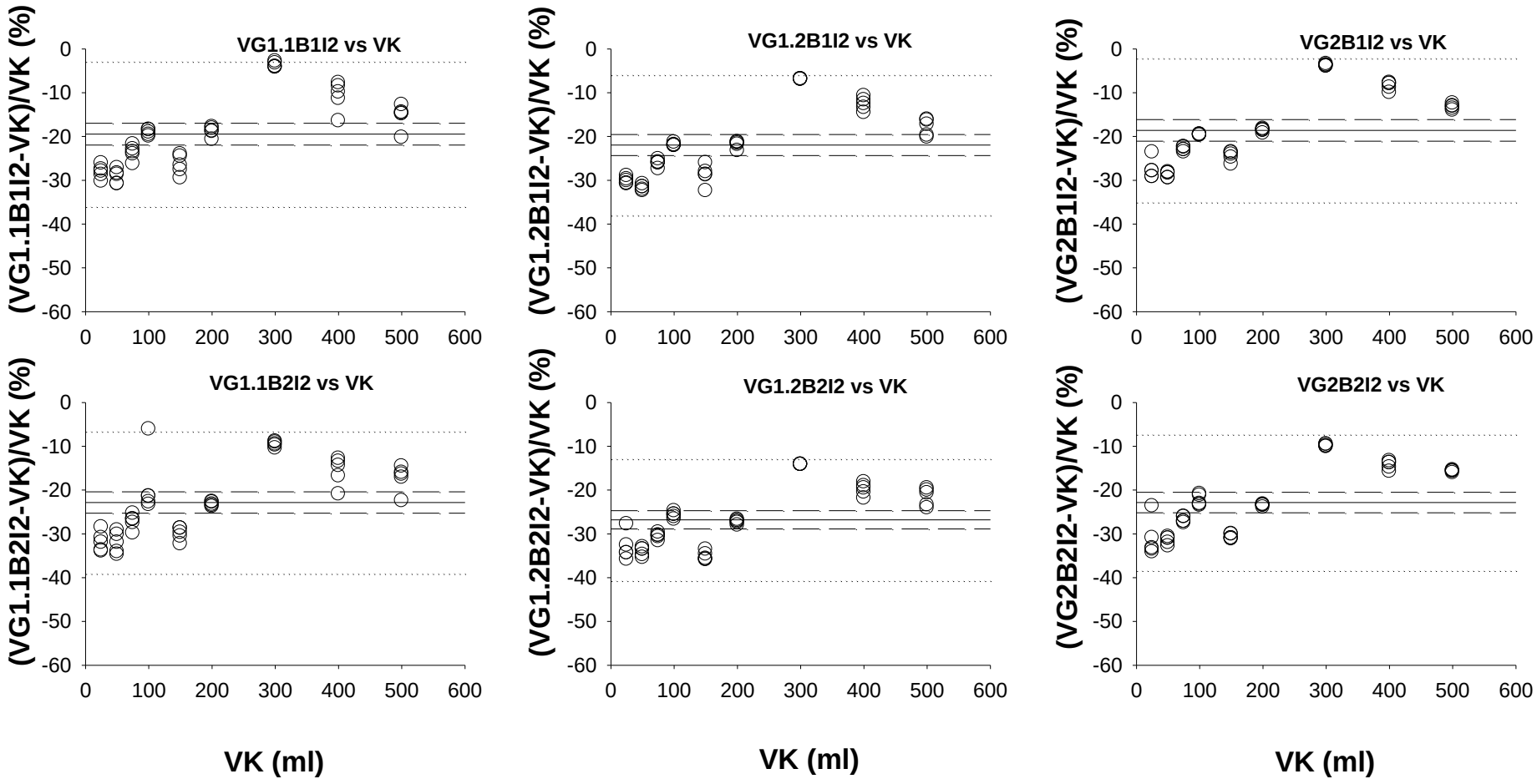


Figure 3-4 Bland Altman plots for 18 algorithms: Comparison between Known Volumes (VK) and Measured Volumes in percentage (%). Solid lines: bias (mean difference); dash lines: 95% confidence interval (CI) of bias; dotted lines: upper and lower limits of agreement

3.3.2. Results of reproducibility assessment

In order to assess the performance of reproducibility, the coefficient of variation (CV) of every five volume measurements obtained by one algorithm on each balloon phantom was calculated, as shown in Figure 3-5. Table 3-3 lists the average ± SD (%) of the CV for all the 18 algorithms in measuring balloon phantoms with nine different volumes. It shows that the averages of the CV for each algorithm (each column) decrease from the first row ($G_{1.1}$) to the last row (G_2).

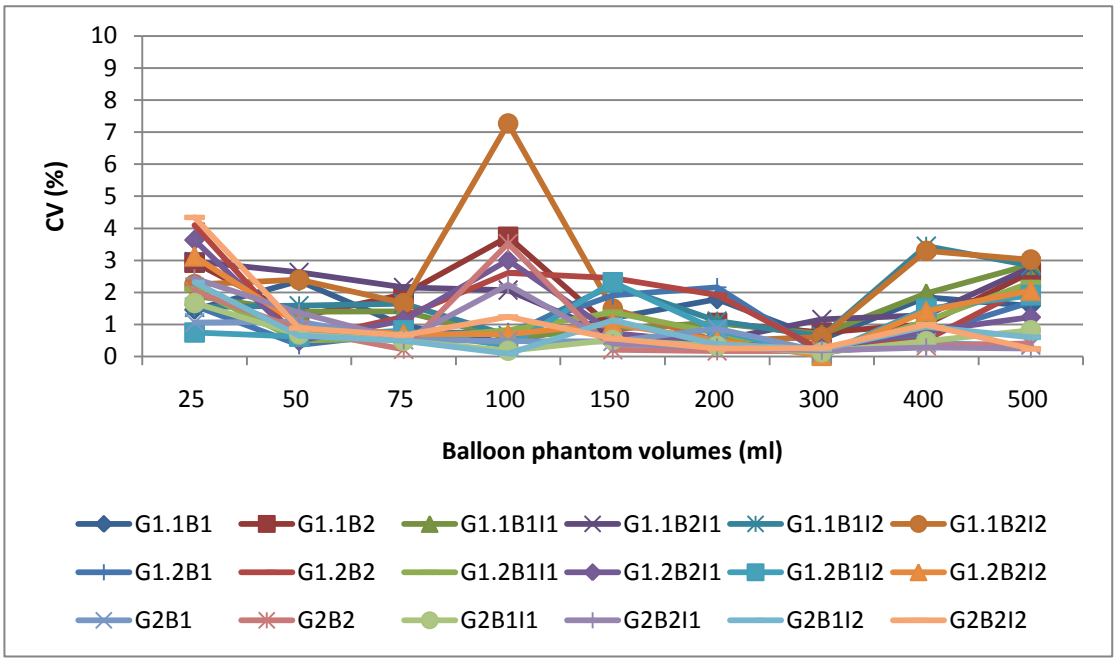


Figure 3-5 Coefficient of variation (CV) of 18 semi-automated algorithms for nine volumes of balloon phantom, showing that the CV was always less than 10%

Table 3-3 Average ± SD (%) of the coefficient of variation (CV) for all the 18 algorithms in measuring balloon phantoms with nine different volumes

	CV: average ± SD (%)					
	B_1	B_2	B_1I_1	B_2I_1	B_1I_2	B_2I_2
$G_{1.1}$	1.4±0.6	1.8±1.1	1.4±0.7	1.8±0.9	1.7±1.0	2.5±2.0
$G_{1.2}$	1.1±0.7	1.7±1.3	1.1±0.7	1.3±1.2	1.0±0.7	1.2±0.9
G_2	0.6±0.3	0.9±1.2	0.6±0.5	0.9±0.9	0.8±0.7	1.0±1.3

After applying ANOVA, significant differences were detected between each two of the three edge operators: $P<0.05$ for $G_{1.1}$ and $G_{1.2}$, $G_{1.1}$ and G_2 as well as $G_{1.2}$ and G_2.

Mean of the CV: $G_{1.1}$= 1.8%; $G_{1.2}$=1.2% and G_2=0.8%. There was also a significant difference between border detection algorithm B_1 and B_2, although the P value was very close to 0.05 (P=0.04; Mean of the CV: B_1=1.1%; B_2=1.5%). No significant difference was found for using or not using interpolation algorithms: P=0.68 for no interpolation and using I_1; P=0.43 for no interpolation and using I_2; P=0.23 for I_1 and I_2. Mean of the CV: no interpolation=1.3%; I_1=1.2%; I_2=1.4%.

Generally, all 18 algorithms showed very good repeatability in the balloon phantom volume measurements. Figure 3-5 demonstrates that the CV of all the algorithms was always smaller than 10%. Some of them were even close to 0, which meant the five volume measurements gave similar results.

The edge operator G_2 provided the best performance, and $G_{1.2}$ was better than $G_{1.1}$. This phenomenon can also be observed in Figure 3-3 and Figure 3-4. In the third column of Bland Altman plots (all using G_2), the group of five measurements of each balloon volume overlapped mostly. In the second ($G_{1.2}$) and first columns ($G_{1.1}$), the measurements overlapped less and less. For border detection algorithm B_1 and B_2, the difference was very small, although statistically significant. There was no significant difference between applying the interpolation algorithms or not, and there was no significant difference for applying I_1 or I_2.

3.3.3. Results of finite element model assessment

After previous statistical analysis, the performance of each algorithm became clear. They all significantly underestimated the balloon phantom volumes. The 3D gradient operator G_2 showed better results of volume measurements than 2D operators in both accuracy and reproducibility assessments. Algorithms with interpolation worked much better than those without interpolation for accuracy. For both accuracy and repeatability, although the results of B_1 and B_2 showed significant differences, after applying interpolation, the differences were not significant, no matter which interpolation (I_1 or I_2) was applied. In order to investigate the performance of the finite element model, the traced geometric data of the phantom inner wall surface from $G_2B_2I_2$ was chosen, because $G_2B_2I_2$ was one of the best algorithms assessed.

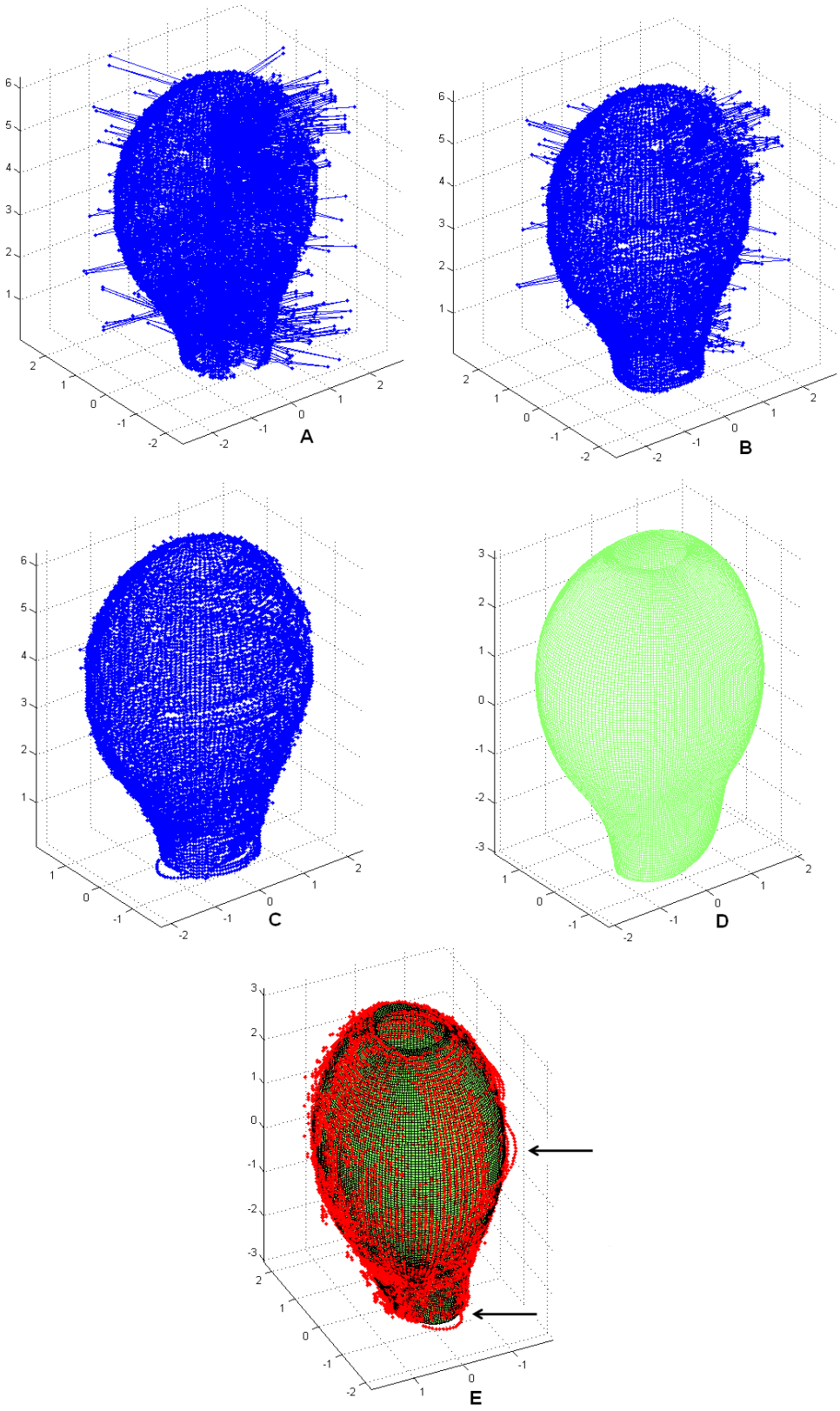


Figure 3-6 3D reconstruction of the 50ml balloon phantom inner wall surface by different algorithms: (A) G_2B_2; (B) $G_2B_2I_1$; (C) $G_2B_2I_2$; (D) FEM from $G_2B_2I_2$; (E) the overlap of $G_2B_2I_2$ (red dots) and FEM (green meshes)

Figure 3-6 demonstrates the 3D reconstruction of the 50ml balloon phantom inner wall surface by four different algorithms (G_2B_2, $G_2B_2I_1$, $G_2B_2I_2$ and FME based on $G_2B_2I_2$) after applying 3D gradient operator G_2. Although the border detection algorithm B_2 succeeded in finding most of the balloon inner wall surface, because of the influences of the noise and artefact in 3D echo images many delineations were obviously incorrect, as shown in Figure 3-6 A. After linear interpolation I_1, the surface tracing became better, especially in the bottom neck area, as shown in Figure 3-6 B. Then, after the elliptical interpolation I_2, Figure 3-6 C showed a smooth surface delineation of the balloon inner wall. However, it was still not perfect. Figure 3-6 D was the finite element model based on the surface tracing data from the $G_2B_2I_2$. If the FEM (green meshes) and the traced surface data (red dots) were overlapped, and then rotated into another orientation, as shown in Figure 3-6 E, it was found that I_2 performed poorly in two slices, which are indicated by the arrows. However, it was improved by the FEM. The phenomenon shown in Figure 3-6 was typical and frequently found during the measurements.

After Bland Altman analysis, it was found that the results of both algorithms showed significant differences with known balloon phantom volumes. The first two plots of Figure 3-7 show that the 95% confidence intervals (dash lines) of both algorithms do not cross the zero axes, which indicated that they both significantly underestimated the balloon phantom volumes. However, the bias±SD of FEM (-27.1±15.8ml) was smaller than $G_2B_2I_2$ (-36.0±21.8ml). Significant differences were also found between FEM and $G_2B_2I_2$ ($P<0.05$, paired Student t test). The 95% confidence interval does not cross the zero axes in the last plot (comparison between $G_2B_2I_2$ and FEM). For reproducibility, as for the previous analysis, the coefficient of variation (CV) of every five measurements of each balloon phantom between $G_2B_2I_2$ and FEM was compared by the Student t test. The result showed no significant difference ($P = 0.40$, Mean of the CV: $G_2B_2I_2$ = 1.0%, FEM = 0.8%). The summary of the comparison between $G_2B_2I_2$ and FEM is also listed in Table 3-4.

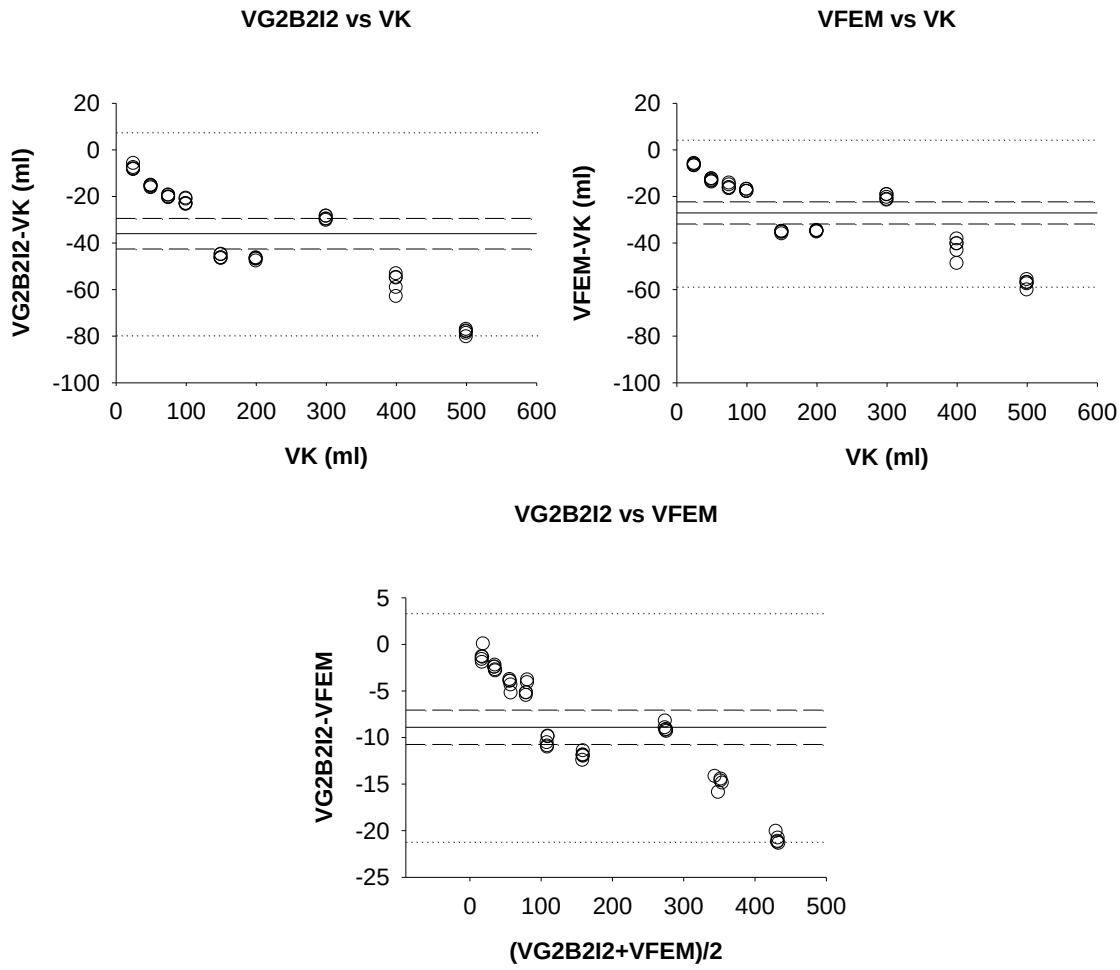


Figure 3-7 Bland Altman plots: First row: comparison between Known Volumes (VK) of balloon phantoms and Measured Volumes (ml) by $G_2B_2I_2$ and FEM separately; Second row: comparison of the Measured Volumes (ml) between $G_2B_2I_2$ and FEM. Solid lines: bias (mean difference); dash lines: 95% confidence interval (CI) of bias; dotted lines: upper and lower limits of agreement

Table 3-4 Comparison between $G_2B_2I_2$ and FEM: average ± SD for the bias with known phantom volumes and the coefficient of variation (CV) of repeat measurements

	Measured - Phantom volume (ml)	CV (%)
$G_2B_2I_2$	-36.0±21.8*	1.0±1.3
FEM	-27.1±15.8*	0.8±0.6

Note: '*' means significant difference.

3.4. Further improvements

After the previous analysis, significant underestimation (at least more that 18% bias) of the balloon phantom volumes was found for all the semi-automated algorithms, including the FEM. This section illustrates why it happened, and how to improve the semi-automated measurement.

3.4.1. What caused the significant underestimation?

There were two possible reasons of causing such significant volume underestimation: the semi-automated algorithms were not tracking the balloon inner wall boundaries correctly on the 3D echo images; or, they did track the correct boundaries, but it was the echo images that did not define the boundaries correctly.

By 3D reconstructing the balloon inner wall surface (Figure 3-6), and highlighting the detected boundary pixels on the 3D echo images slice by slice (Figure 3-9, a and c), it proved that the semi-automated delineation of the boundaries was correct in most of cases after applying the interpolation algorithms (I_1 or I_2). Although a few of incorrect detections existed, it was highly unlikely that they were the main source of causing the more than 18% bias. Therefore, the first possible reason was ruled out. However, can the second reason explain the volume underestimation? To answer that, the imaging resolution was investigated.

3.4.2. Poor definition of the balloon wall boundary

Echocardiography is an imaging modality based on acoustic techniques. The physical property of the ultrasound determines that echocardiography can offer images with excellent temporal resolution, but relatively poorer spatial resolution. Because of the problem of low spatial resolution, the definition of an edge in an echo image is normally unsharp and spread, as shown in Figure 3-8. For an ideal definition of an edge in a digital image, the change of intensity happens in one pixel (Figure 3-8, upper graph). However, in real cases, the intensity variation always happens in several conjunctive pixels (Figure 3-8, lower graph) because of the limitation of the spatial resolution.

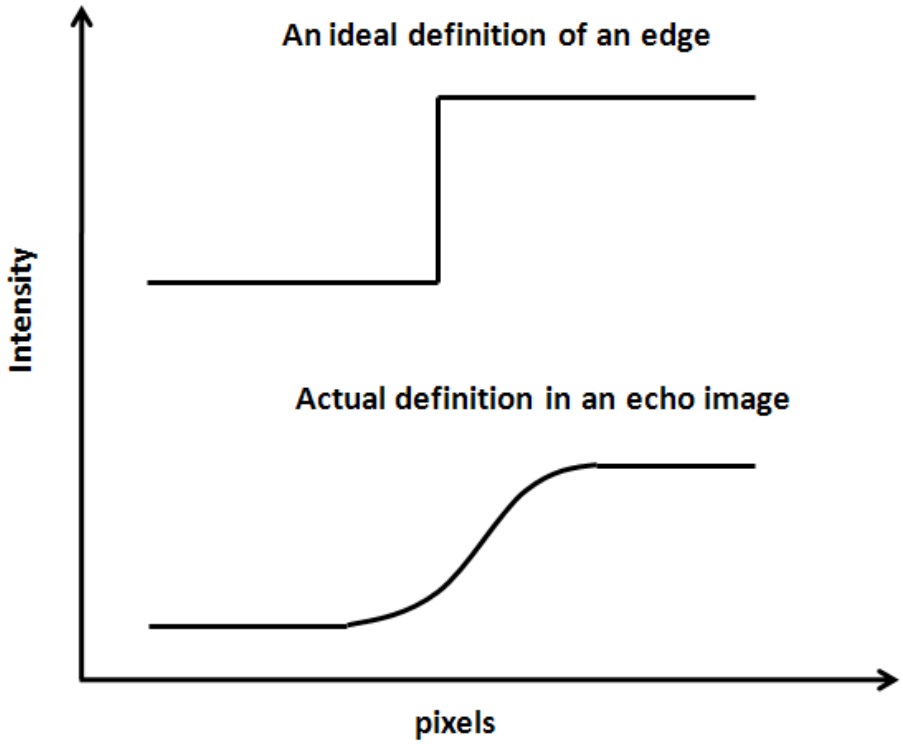


Figure 3-8 Difference between an ideal definition and an actual definition of the same edge in an echo image: upper graph: an ideal definition of an edge, the intensity changes within one pixel; lower graph: an actual definition of an edge, the intensity changes in several pixels.

Figure 3-9 shows an example of two short axis slices extracted from two 3D echo images. Picture (a) was the 25ml balloon (the smallest balloon volume), and (c) was the 500ml balloon (the biggest balloon volume). The delineation (white dots) of the inner wall boundary was done by the algorithm $G_2B_2I_2$. Both picture (a) and (c) shows that the semi-automated delineation was reasonably good. If the delineation was performed manually, the result would be similar to this.

After magnifying the highlighted area (the black square) in (a) and (c), the details of the images are clearly demonstrated in (b) and (d). Picture (b) shows that the balloon wall thickness is normally more than five pixels (The highlighted place is 7 pixels). This was similar for different 3D images of different balloon volumes. Even for the biggest balloon volume (500ml), the definition of the balloon wall thickness in the echo images was normally more than five pixels, as shown in picture (d).

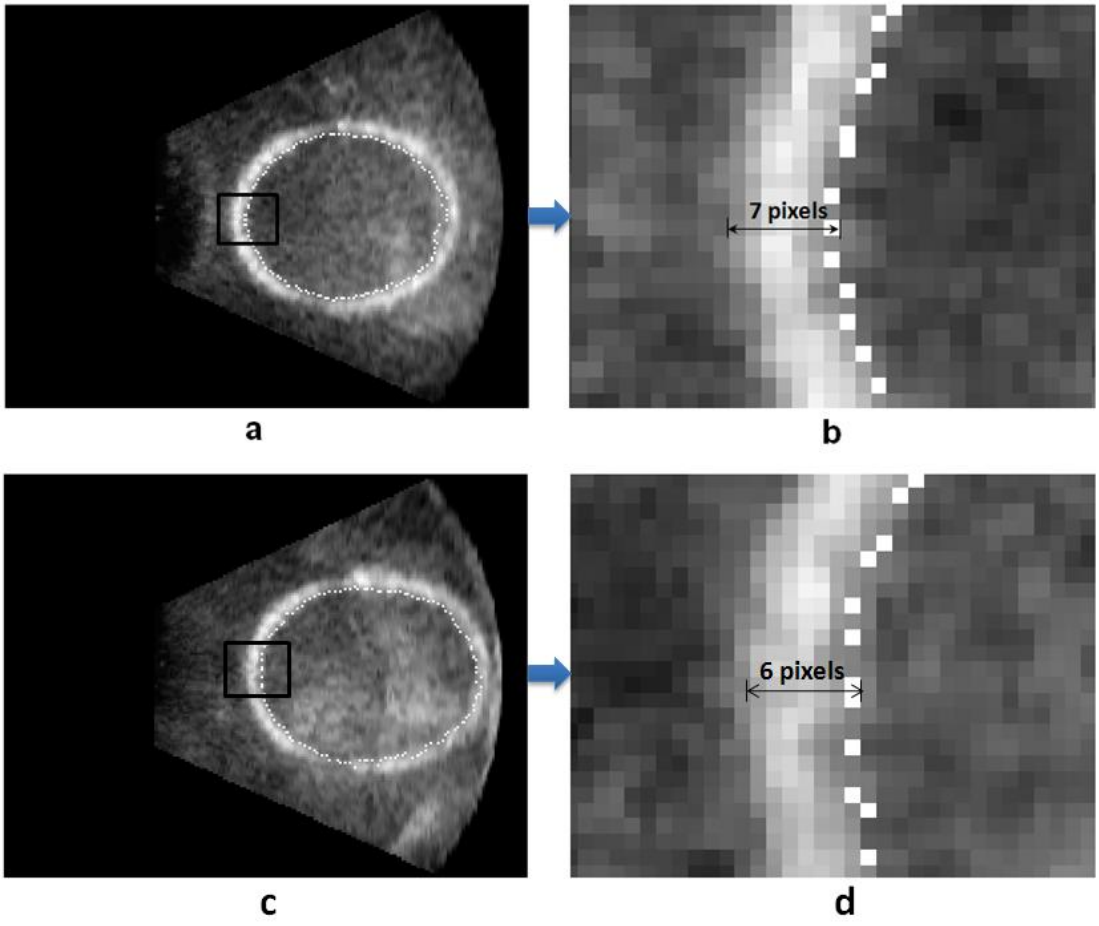


Figure 3-9 Two short axis slice extracted from the 3D echo images of 25ml and 500ml balloon phantom: (a) & (c) Semi-automated delineation of the 25ml and 500ml inner wall boundaries (white dots) by $G_2B_2I_2$; (b) & (d) Magnifying the highlighted black square areas in (a) & (c) respectively

Table 3-5 Variation of voxel sizes of 3D echo images for different balloon phantom volumes

	Balloon Phantom Volumes								
Voxel Size	**25ml**	**50ml**	**75ml**	**100ml**	**150ml**	**200ml**	**300ml**	**400ml**	**500ml**
x (mm)	0.06	0.06	0.08	0.08	0.10	0.10	0.11	0.12	0.12
y (mm)	0.06	0.06	0.08	0.08	0.10	0.10	0.12	0.12	0.12
z (mm)	0.05	0.05	0.06	0.07	0.08	0.08	0.09	0.10	0.10

In this study, all the 3D images of different balloon phantoms had the same number of voxels (x × y × z: 144 × 160 × 208), but the actual voxel size varied from

0.06mm×0.06mm×0.05mm (25ml balloon) to 0.12mm×0.12mm×0.10mm (500ml balloon). Table 3-5 lists the variation of voxel sizes of different 3D echo images. It indicates that the voxels were getting bigger while the balloon volume increased.

For the 25ml balloon phantom, the short axis plan was the x-y plan. Therefore, the wall thickness was more than 0.3mm (0.06mm×5) in an image. For the 500ml balloon phantom, the wall thickness was more than 0.6mm (0.12mm×5). This did not fit the actual balloon wall thickness, which was less than 0.1mm. Moreover, the actual balloon wall thickness should be thinner instead of being thicker, as the balloon volume increases.

It was the low spatial resolution of the echo image that caused the spread effect (as shown in Figure 3-8) of the balloon wall boundary that broadened the wall thickness. This resulted in poor definition of the balloon inner and outer wall boundaries.

For the semi-automated algorithms, they traced the edges of the inner wall boundary in the echo images, as shown in Figure 3-9, (a) and (c) (white dots). However, because the balloon wall was broadened in the images, the detected edges were not the right locations of the real balloon inner wall boundaries, but the broadened boundaries.

The real inner wall boundary was likely to locate inside the balloon wall in the echo images. Therefore, this poor definition of the balloon wall boundary caused a certain underestimation of the balloon phantom volumes. Then, the question was how big an underestimation was caused by this?

3.4.3. Find correct boundaries on 3D echo images

In order to investigate the location of the actual balloon inner wall boundary, and also verify the hypothesis for the underestimation of the balloon phantom volumes, another analysis was performed.

Four semi-automated algorithms were applied to measure the nine balloon phantom volumes, which were $G_2B_1I_1$, $G_2B_2I_1$, $G_2B_1I_2$ and $G_2B_2I_2$. Based on the previous analysis, they were the four best algorithms in accuracy assessment, and they did not give significant differences in volume measurements.

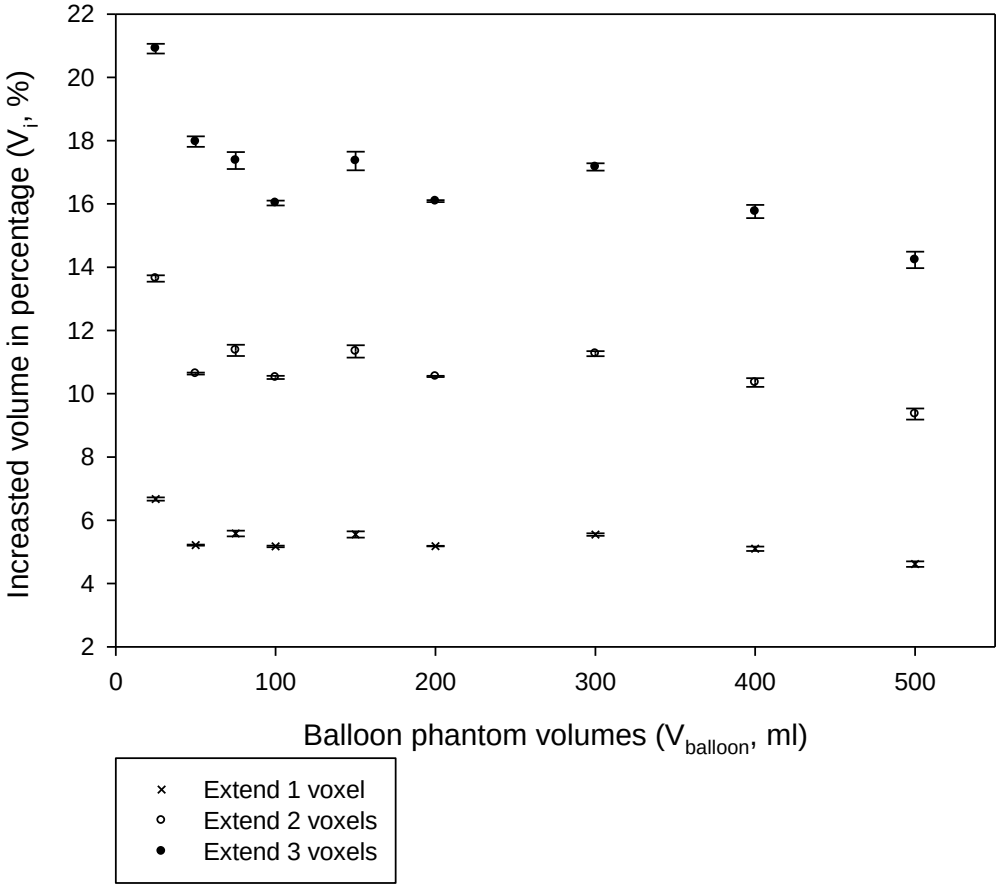


Figure 3-10 A plot of the increase of volume measurements in percentage (Average and SD of V_i of the four algorithms: $G_2B_1I_1$, $G_2B_2I_1$, $G_2B_1I_2$ and $G_2B_2I_2$) by extending the detected inner wall boundaries towards outside for one, two and three voxels.

For each balloon phantom and each algorithm, the delineation of the balloon inner wall surface was performed four times with the same ROI. In the first time, each algorithm used the default setting. In the second time, all the detected edges (inner wall boundaries) were extended outwards for one voxel. In the third and fourth times, the detections were extended outwards for two and three voxels. After that, for each balloon volume ($V_{balloon}$), the measured volume was coded as V_{xy}, where x = 1, 2⋯4, represented the four algorithms; y = 1, 2⋯4, represented the four measurements. For example, V_{12} meant that the second measurement was extended one voxel, compared with the measurement obtained by the first algorithm $G_2B_1I_1$. Then, for each balloon volume, the increased volumes V_i (in percentage) caused by extending the detected edges outwards for one, two and three voxels were calculated separately:

$$V_i = (V_{x(n+1)} - V_{x1})/V_{balloon}$$

Where x = 1, 2⋯4 (four algorithms); n = 1, 2 and 3 (extending 1, 2 and 3 voxels). Finally, for each extension of the detected edges, the average and the standard deviation of V_i of the four algorithms were calculated and demonstrated against the balloon volumes, as shown in Figure 3-10 and Table 3-6.

Table 3-6 Increase of volume measurements in percentage (Average and SD of V_i of the four algorithms: $G_2B_1I_1$, $G_2B_2I_1$, $G_2B_1I_2$ and $G_2B_2I_2$) by extending the detected inner wall boundaries towards outside for one, two and three voxels

	Extend 1 voxel		Extend 2 voxels		Extend 3 voxels	
Balloon volume (ml)	Average (%)	SD (%)	Average (%)	SD (%)	Average (%)	SD (%)
25	6.67	0.05	13.64	0.10	20.91	0.15
50	5.22	0.02	10.64	0.03	17.97	0.17
75	5.58	0.09	11.37	0.18	17.37	0.27
100	5.17	0.03	10.52	0.05	16.03	0.08
150	5.55	0.10	11.34	0.20	17.36	0.29
200	5.18	0.01	10.55	0.02	16.09	0.02
300	5.55	0.04	10.27	0.08	16.17	0.12
400	5.10	0.07	10.35	0.14	15.76	0.21
500	4.61	0.09	9.36	0.17	14.23	0.26

Figure 3-10 and Table 3-6 illustrates that after extending the detected inner wall boundaries outwards by three pixels, the measured volumes gained at least extra 14% of the balloon phantom volume. This directly explained the reason why the balloon phantom volumes were underestimated by an average 20%. It also indicated that the true balloon inner wall boundaries were likely to locate at the place three voxels inside the balloon inner wall edges of the echo images.

Table 3-7 lists the final absolute and relative results of volume measurements given by the four algorithms ($G_2B_1I_1$, $G_2B_2I_1$, $G_2B_1I_2$ and $G_2B_2I_2$) after extending the detected inner wall boundaries towards outside for three voxels). Obviously, the biases with known balloon volumes were much smaller.

Table 3-7 Absolute and relative results of volume measurements given by the four algorithms ($G_2B_1I_1$, $G_2B_2I_1$, $G_2B_1I_2$ and $G_2B_2I_2$) after extending the detected inner wall boundaries towards outside for three voxels

Balloon volume (ml)	Absolute values (ml)				Relative values (%)			
	$G_2B_1I_1$	$G_2B_2I_1$	$G_2B_1I_2$	$G_2B_2I_2$	$G_2B_1I_1$	$G_2B_2I_1$	$G_2B_1I_2$	$G_2B_2I_2$
25	24.4	24.0	24.3	24.3	97.6	96.2	97.2	97.2
50	44.7	44.7	44.7	43.4	89.5	89.4	89.5	86.8
75	72.9	73.5	71.0	68.3	97.2	97.9	94.7	91.1
100	93.8	91.6	96.5	92.9	93.8	91.6	96.5	92.9
150	142.5	135.2	140.1	130.6	95.0	90.1	93.4	87.1
200	195.3	187.5	195.0	185.3	97.6	93.8	97.5	92.6
300	320.0	316.8	327.0	320.0	106.7	105.6	109.0	106.7
400	396.4	390.7	398.1	395.8	99.1	97.7	99.5	98.9
500	489.7	488.9	495.6	494.5	97.9	97.8	99.1	98.9

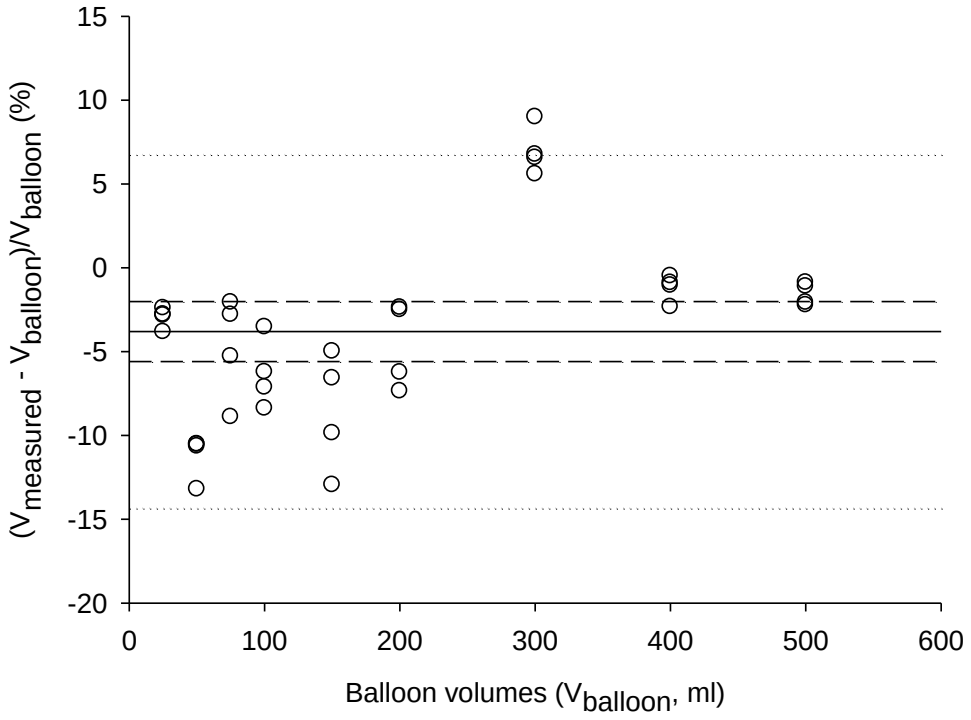


Figure 3-11 Bland Altman plot of comparing balloon volumes ($V_{balloon}$) with measured volumes ($V_{measured}$) given by the four algorithms after extending three voxels of the inner wall edges; The comparison is in %: ($V_{measured} - V_{balloon}$)/ $V_{balloon}$; Solid lines: bias (mean difference); dash lines: 95% confidence interval (CI) of bias; dotted lines: upper and lower limits of agreement

After that, the Bland Altman method was applied to the measured volumes given by the four algorithms after extending three voxels of the inner wall boundary. The comparison was shown in Figure 3-11. The result shows that the average bias was only -3.8%, although statistically significant difference still existed (the 95% CI did not cross the zero axis).

Both Table 3-7 and Figure 3-11 indicate that after extending the detected wall edges, the four algorithms showed an overestimation of the balloon phantom volume only in measuring the phantom with known volume of 300ml. This was probably an outlier. The most likely reason caused that might be the inaccurate preparation of the balloon phantom. Maybe more than 300ml water was filled into the balloon during the imaging acquisition process. But anyway, this outlier did not affect the overall conclusion of this study.

3.5. Discussion

The quantitative volume determination of heart cavities through imaging technologies can provide important implications for patients with cardiovascular diseases. Despite much research in this area, and the large volume of data published on this topic, there is currently no optimal method.

One particular problem of using echocardiography is the negative influence of high noise and artefact that is inherent to acoustic signals. Although people have proposed 3DE to overcome some limitations of 2DE, this inherent problem yet remains unsolved, and it creates challenges for developing accurate and reliable semi-automated quantification techniques.

By designing and performing the experiment of using the laboratory balloon phantoms to evaluation the performance of the new semi-automated algorithms in border tracing and volume determination, the results proved that 3D echocardiography can decrease the influence of noise and artefact if the 3D echo image is analyzed in an appropriate way. This section discusses why 3D imaging should be better.

Besides the influence of noise and artefact, another inherent limitation of echocardiography is the low spatial resolution. In this study, even with interpolation and FEM, semi-automated algorithms still showed about 20% underestimation of the balloon phantom volumes. Section 3.4 indicates that the main reason was related to imaging resolution. The actual balloon inner wall boundaries were likely to locate at the place three voxels inside the balloon inner wall edges of the echo images. However, after this improvement, 3.8% volume underestimation still existed. This section discusses what caused this small bias.

3.5.1. Why 3D analysis was better?

For ultrasound images, the signal to noise ratio is often poor, and acoustic noise can confound the detection of the cardiac chamber wall boundary. This inherent limitation is the biggest problem of automatic or semi-automatic delineation of the cardiac chamber wall surfaces.

After applying edge detection operators, the problem can be partially alleviated by extending the size of the neighbourhoods over which the differential gradients are computed (64), but the apparent width of the edges (wall boundaries) will be broadened as a consequence, which leads to a systematic underestimation of the chamber volumes. Therefore, there is a trade-off between noise reduction and resolving power.

With 3D edge operators, the additional third direction for the operator can gather extra information from the neighbourhood whilst keeping the size in all directions relatively small. For this reason, the 3D operators make full use of the advantage of 3D echo images to limit the artefacts and noise effects on the images without further broadening the edges of the boundaries.

Figure 3-12 illustrates a typical example. In the highlighted area, there is some missing information along the balloon wall in the original image (a). After applying 5×5 and 11×11 2D edge operators, the gaps along the wall still exist in both gradient image (b) and (c). The major difference of (b) and (c) is that the double layer wall boundary is thicker in (c), which proved that increasing the operator size broadens the edges. However, the gaps are restored in image (d) because of the action of the $5\times5\times5$ 3D

operator, and the thickness of the edges is more or less the same as it is in image (b). This explained why the 3D edge operator showed the best performance of volume measurements in both accuracy and reproducibility.

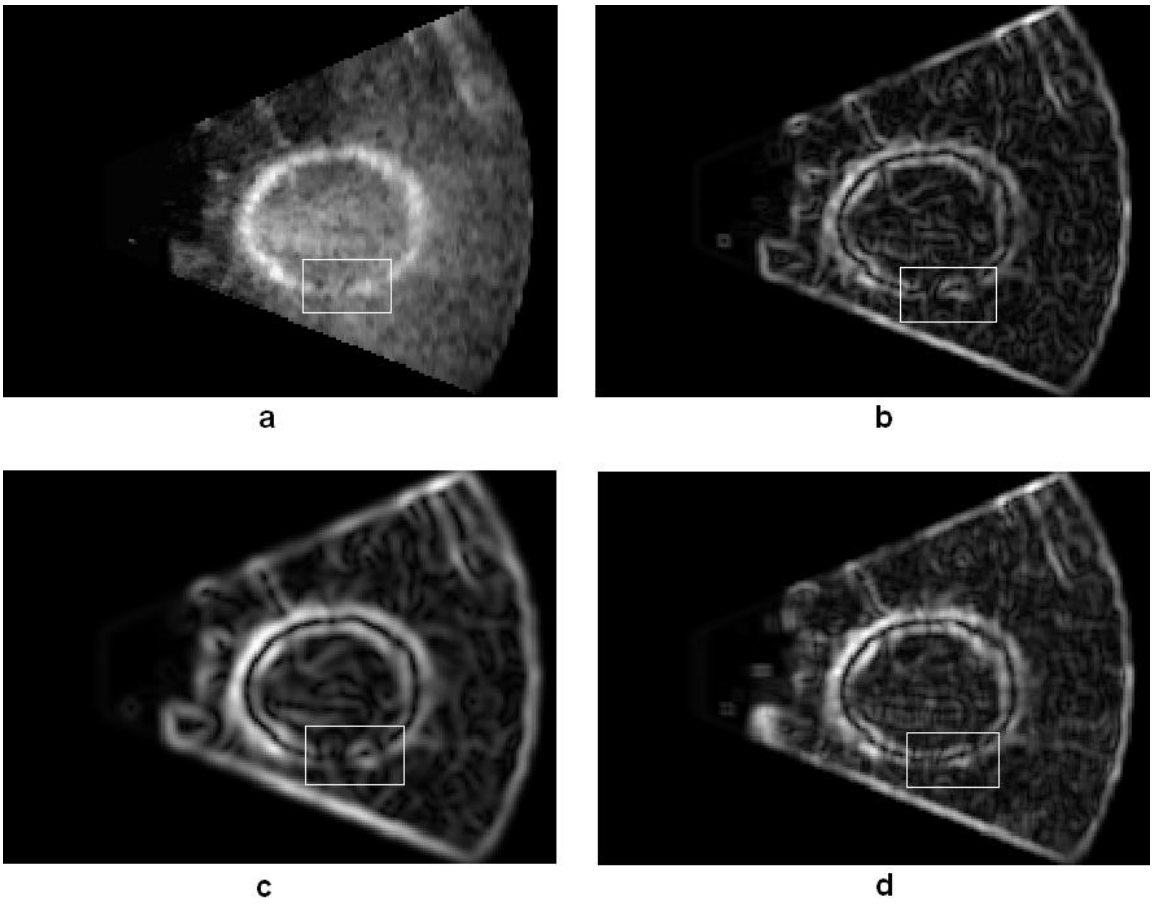


Figure 3-12 Effects of 2D and 3D edge operators: (a): a 2D short axis slice extracted from a 3D image of a balloon image; (b): gradient image obtained by using 5×5 Macleod operator; (c): gradient image obtained by using 11×11 Macleod operator; (d): gradient image obtained by using 5×5×5 Macleod operator

Although the influence of noise and artefact on echo images can be reduced by 3D edge operators, the problem is not able to be solved completely. In Section 3.3.3, Figure 3-6 (a) demonstrates that after applying the 3D operator G_2, the automated border detection algorithm B_2 still suffered from the poor quality of the echo images. Therefore, border interpolation algorithms (I_1 or I_2) improved the border delineations by adding geometric assumptions (linear or elliptical). However, either I_1 or I_2 has its limitations.

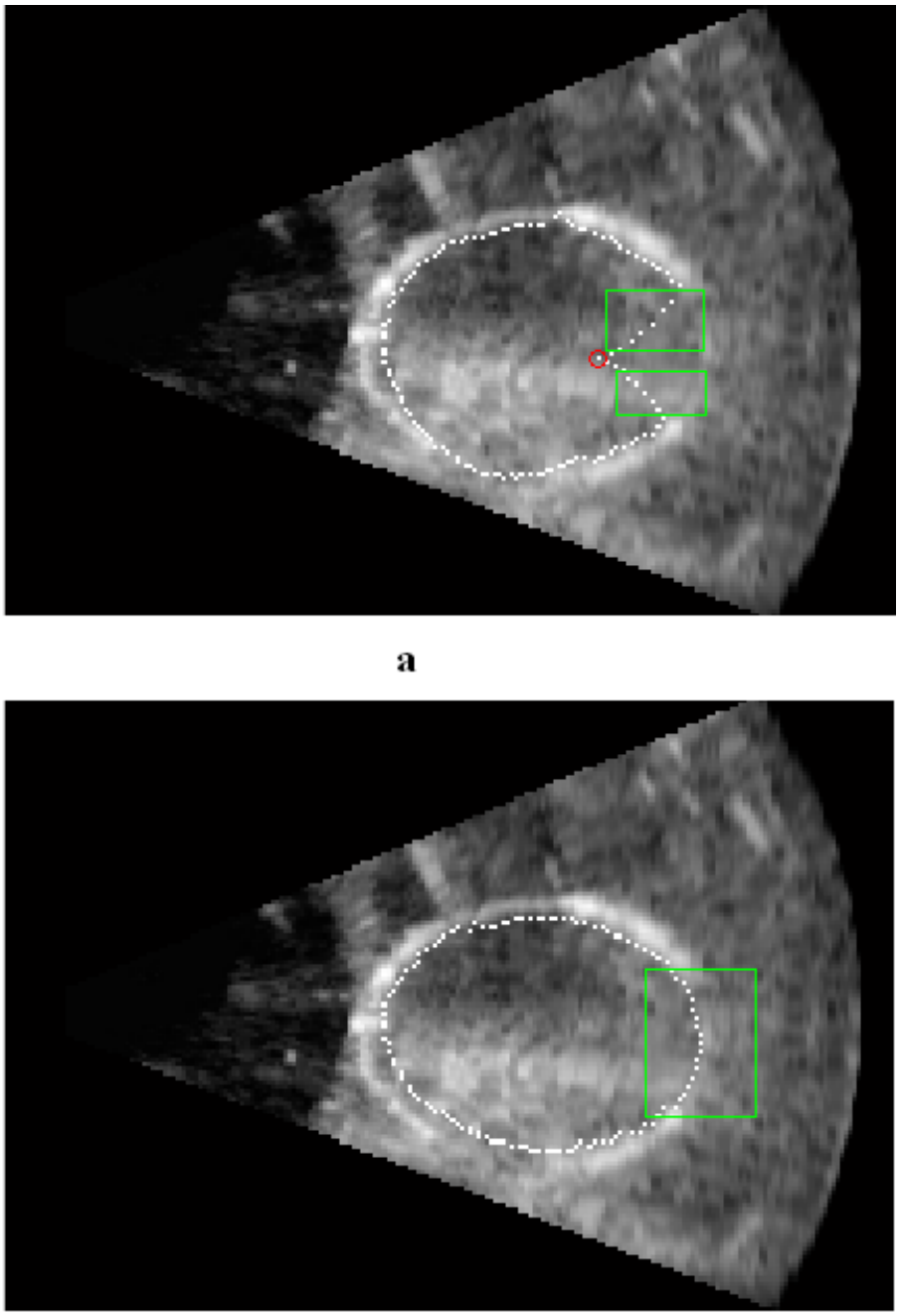
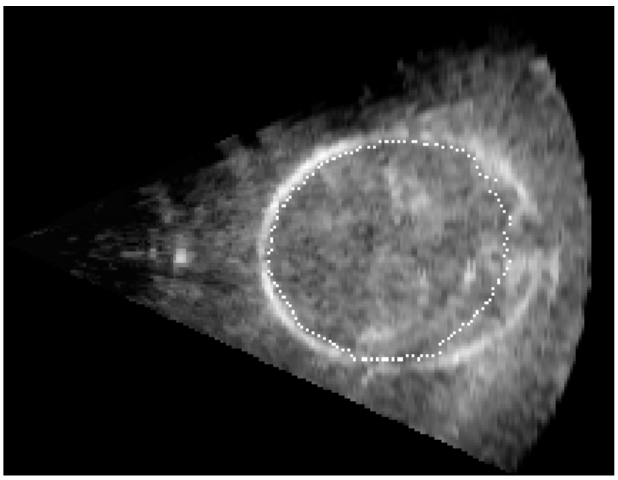


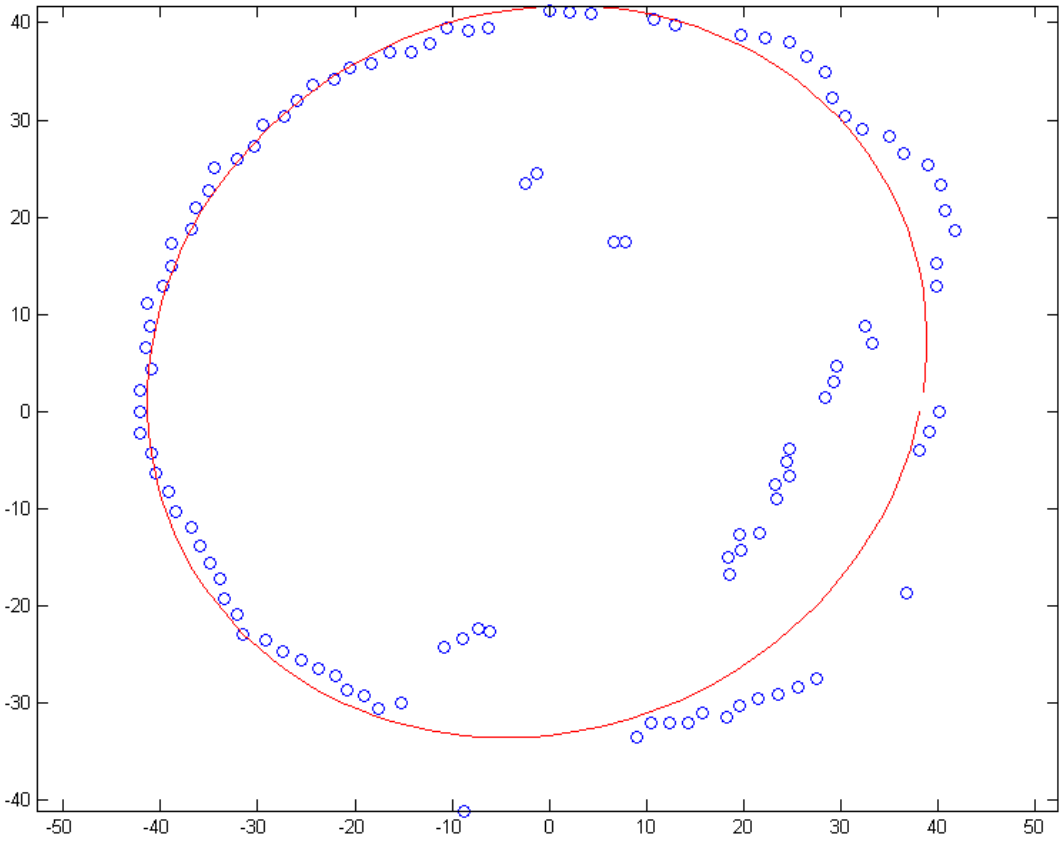
Figure 3-13 Comparison of linear and elliptical interpolation: the manual initializations of both pictures are the same. White dots are the delineation of inner wall borders. (a) Linear interpolation I_1; (b) Elliptical interpolation I_2

Figure 3-13 shows an example when the linear interpolation failed but elliptical interpolation succeeded in tracing the inner wall border. This is because I_1 only used the two detected points neighbouring undetected points to interpolate (Chapter 2, Section 2.4). If one of the two detected points was not the true border, as in the case in Figure

3-13 (a) where the detected border (white dot) inside the red circle was actually a detection of the noise, the linear interpolation inside the green squares failed.



a



b

Figure 3-14 An example of poor performance of elliptical interpolation: (a) Phantom inner wall tracing (white dots) by algorithm $G_2B_2I_2$; (b) Fitted ellipse (red ellipse) based on the detected inner wall borders (blue circles)

Different from I_1, I_2 used all the detected borders instead of only two borders to fit an ellipse (Chapter 2, Section 2.4), which can minimize the influence from one or several incorrect detections. Therefore, Figure 3-13 (b) shows better border delineation. It should be pointed out that the interpolated white dots inside the green square in Figure 3-13 (b) are not very smooth. This was because of the low imaging resolution, which cannot offer a smooth curve of the interpolated boundary.

Although I_2 showed certain advantages, I_2 was not able to work well in some circumstances. Figure 3-14 is one example. In the image (a), severe noise appeared in a large area, resulting in tracing of noise rather than the true boundaries, as shown in image (b), blue circles. Because of that, the fitted ellipse (red) still offered an incorrect interpolation.

To overcome this problem, the Finite Element Model (FEM) was applied. The philosophy behind this is the same as applying the 3D gradient operators: The analysis of the 3D echo images should be performed in all three dimensions. Instead only using the traced borders from one 2D plane to interpolate the unsuccessfully traced borders in the same plane, the FEM fits a 3D surface onto the traced borders from all three dimensions. Even if the severe noise appears in one or several planes, the FEM might be able to get the useful information from the other planes to alleviate the effects. This explained the reason why in Figure 3-6, from the image A to D, the balloon phantom inner wall reconstruction improved.

3.5.2. Why was there 3.8% underestimation of volumes?

After investigating the echo image resolution and seeking the true balloon inner wall boundaries on those images, the bias of semi-automated volume measurements were generally improved from about 20% to only 3.8%. However, the differences with the actual phantom volumes were still significant. This was because that there was a flaw of the ROI establishment.

The manual initialization of a ROI of a 3D balloon image needed two short axis slices (Chapter 2): the top and bottom slices, as shown in Figure 3-15, the two blue slices. In each of them, the balloon wall contour must enclose a reasonably big area so that the

manual selection of the ‘central point’ inside the wall contour can be carried out easily. This caused two parts of the balloon to be excluded from the ROI, which were the parts above the top slice and below the bottom slice. Therefore, it was impossible to detect them by any of the semi-automated algorithms, and so these two parts were excluded from the balloon volume. For this reason, even though an algorithm was able to delineate the balloon inner wall surface perfectly correctly, it could only measure the balloon volume inside the ROI, and so give an unavoidable underestimation of the phantom volume.

This flaw existed for all 18 semi-automated algorithms and the FEM too. Figure 3-6 shows that the 3D reconstruction of the balloon inner wall surface did not cover the whole balloon phantom. There were some parts missing on the top and the bottom of the balloon, which illustrates this problem. Therefore, these two missed parts explained the average 3.8% underestimation of the balloon volumes.

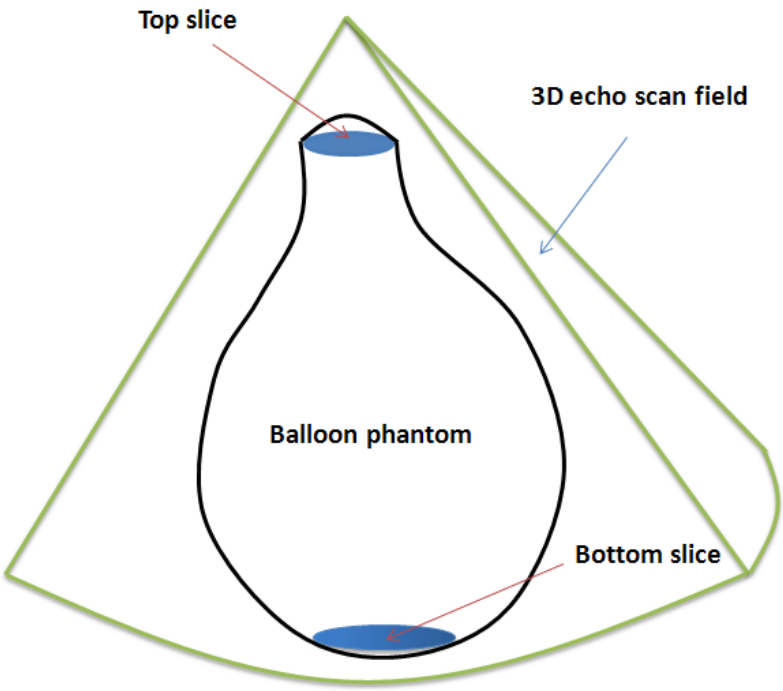


Figure 3-15 Demonstration of the volume underestimation caused by the selection of two slices to build up the ROI: the top and bottom parts of the balloon were excluded from the ROI

3.6. Conclusion

The semi-automated algorithms were able to trace the balloon phantom inner wall surface and measure volumes quantitatively from 3D echocardiography images. The algorithms were easy to set up and were user friendly, as only two points were required to be selected manually, and the scan radius R needed to be pre-defined to establish the ROI. However, they all showed significant underestimation of the phantom volumes.

The 5×5×5 3D edge operator showed the best performance in both accuracy and reproducibility in measuring the balloon phantom volumes. The 5×5 2D operator showed better performance in accuracy than the 11×11 operator did, however, the 11×11 operator showed better performance in reproducibility than the 5×5 operator did.

The interpolation algorithms significantly improved the volume measurements in accuracy. For both accuracy and repeatability, although the results of B_1 and B_2 showed significant differences, after applying interpolations the differences became non-significant, no matter which interpolation (I_1 or I_2) was applied.

Comparing with $G_2B_2I_2$, which was one of the best algorithms, the finite element model (FEM) further improved the volume measurements in accuracy. However, no significant difference was found in reproducibility.

The poor boundary definition caused by the limitation of 3D echo image spatial resolution resulted in a significant underestimation of the balloon phantom volumes. The measurements was improved from about 20% bias to 3.8% bias by extending the detected edges outwards for three voxels.

To sum up this study, it was found that: In order to achieve a better quantitative measurement of 3D echocardiography images, the imaging analysis should be performed from all three dimensions, as this utilized the information contained in 3D images, rather than split them into multiple 2D slices and analyzed them separately. Analyzing 3D echo images in 3D can minimize the negative influence of noise and artefact, and so improve the accuracy as well as reproducibility in quantitative measurement.

Besides noise and artefact, low spatial resolution of 3D echo images also affected the phantom volume measurements. To compensate this influence, true boundaries of balloon phantom inner wall surfaces should be three voxels outwards the detected edges on 3D echo images.

In the next chapter, a new study is introduced. In that study, better tissue-mimicking phantoms were designed, which had the same ultrasound properties as human soft tissue and blood. The differences in quantitative measurement between 2D and 3D echocardiography were investigated.

Chapter 4. Quantitative volume measurements of ultrasonically tissue-mimicking phantoms

In Chapter 3, the newly developed semi-automated surface tracing and volume calculation algorithms were applied to 3D echo images of laboratory balloon phantoms with known volumes. Surfaces of phantom inner walls were traced, and phantom volumes were calculated quantitatively. By comparing calculated volumes with actual volumes, the accuracy as well as the reproducibility of each algorithm was assessed. However, there were some inherent drawbacks of previous imaging acquisition, which limited the reliability of the algorithm assessments.

The first drawback was the ultrasound propagation speed is different in different mediums. The difference of the ultrasound speed in water and in human soft tissue or blood caused certain bias of the volume calculation. Therefore, a theoretical correction was applied to compensate for this difference (see section 3.2.2).

The second drawback was the wall thickness of the balloon phantoms. The balloon wall is much thinner than any part of normal human myocardial wall. Its material (rubber) has totally different ultrasound characteristics with human myocardium. Therefore, it was not an ideal simulation of human heart.

The last drawback was the volume control. The balloon phantom was filled with certain amount of water manually. This process already caused certain bias between its real volume and the expected volume.

In this chapter, new tissue-mimicking phantoms were made. New imaging acquisition system was designed. Special controlled saline water was used instead of ordinary water. All these efforts were made to ensure a precise control of the reference volumes as well as an ideal simulation of human hearts. The new algorithms were applied on both conventional 2D images and 3D images, which were acquired by an updated new Philips 3D ultrasound system. The purpose of this chapter is to find out the strengths and weakness of 2D and 3D echo images, and to assess the performance of the new semi-automated algorithms, so the findings of Chapter 3 can be testified or overthrown.

4.1. Imaging acquisition

The fundamental problems of the previous imaging acquisition were that: the volumes of the balloon phantoms were not precisely controlled; the balloon phantoms and water do not have similar ultrasound properties as human hearts. In order to overcome these problems, new phantoms were made by casting tissue-mimicking materials from industrial lathe cut moulds, and special controlled saline water was used instead of pure water.

4.1.1. Production of the tissue-mimicking material for ultrasound

In 1998, Ernest L. Madsen etc. published a paper in Ultrasound in Med. & Biol. (68), which described a method of production of tissue-mimicking material for ultrasound. It exhibits the 1540 m/s propagation speed and proportionality of attenuation coefficient and frequency over the diagnostic range. This method was chosen for making new tissue-mimicking phantoms, and the following is the step-by-step procedure for producing 1 litre of the material with an attenuation coefficient slope of 0.5 dB cm^{-1} MHz^{-1} and propagation speed of 1540 m/s.

(1) Dissolve 1g of thimerosal into 500 ml of evaporated milk in a 500 ml beaker.

(2) Raise the temperature of the mixture to 68 ℃ in a double boiler, stirring sufficiently to prevent the temperature of any part of the mixture from reaching 72 ℃. Cool the mixture to 55 ℃. This process can be done simultaneously with step (5).

(3) Mix 322ml distilled water and 28 ml of n-propanol at room temperature in a 500 ml beaker.

(4) Rapidly mix in 14 g of dry high-purity agarose (Type 1-A, low EEO, 0.09-0.13 agarose) into the propanol beaker.

(5) Cover the beaker with a polymer food wrap and warm in a double boiler to 90 ℃ until the mixture is transparent. (A puncture in the polymer food wrap will ensure no undesirable pressure difference while still suppressing evaporation of the beaker contents.) The two 500 ml beakers can be heated (one with evaporated milk and

thimerosal, and the other with distilled water and agarose) together by the electrical boiler.

(6) Cool the molten agarose solution to 55 ℃ by immersing the lower part of the beaker in a large container of cold water, stirring the agarose to ensure that no part of it is cooled below 37 ℃ (congealing point).

(7) Combine the 500 ml of 55 ℃ evaporated milk plus thimerosal with 500 ml of 55 ℃ molten agarose and mix well.

(8) After cooling to about 45 ℃, the liquid can be poured into moulds and allowed to cool. Gelling occurs at 36±1 ℃. The melting point of the resulting solid gel is above 80 ℃.

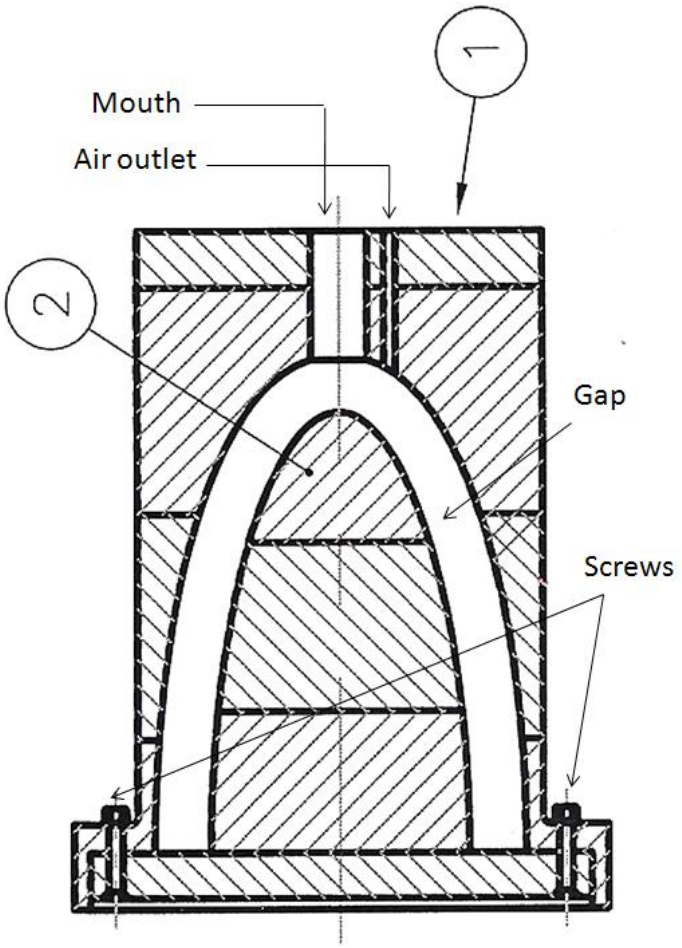


Figure 4-1 Blueprint of the cross section of two mould pieces: 1 is the outer piece and 2 is the inner piece. They are screwed together.

4.1.2. Mould design and production

A semiellipsoid (semi-long axis) is frequently used to describe a simplified geometry of human LV. Therefore, the new moulds were designed to make the phantoms having precise semiellipsoidal shape. Figure 4-1 shows the design of the two mould pieces: outer Piece 1 and inner Piece 2. They were screwed together tightly to ensure there was no leakage. The tissue-mimicking material was poured from the mouth of the Piece 1 into the gap space between the two mould pieces. A small hole was drilled right next to the mouth of Piece 1 (air outlet), so that the air can come out during the pouring process without producing any air bubbles.

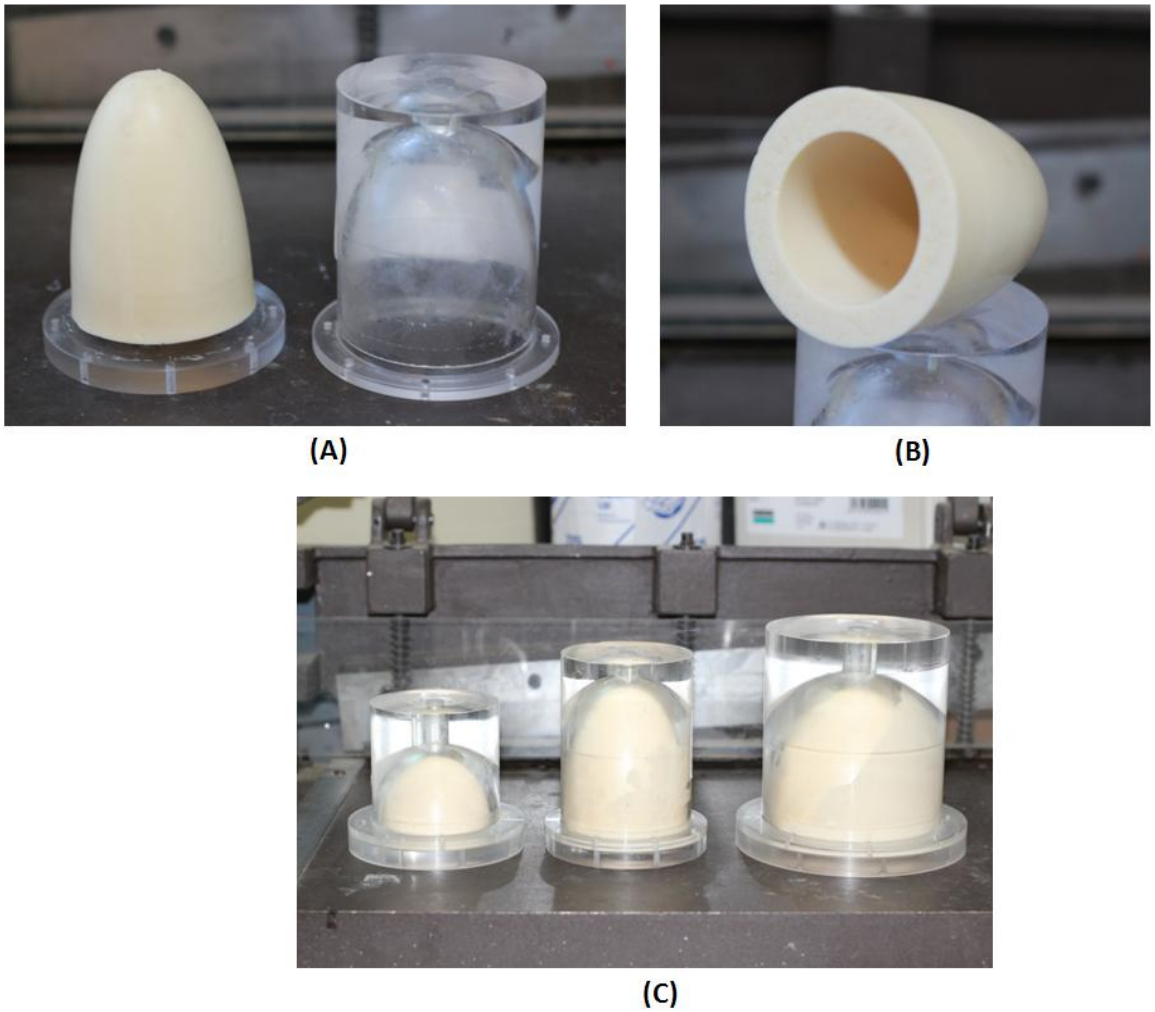


Figure 4-2 Pictures of tissue-mimicking phantoms: (A) moulds were opened after the solidification of the phantom; (B) A final finished phantom with a perfect semiellipsoidal cavity; (C) Three phantoms were made in three different volumes

After pouring the tissue-mimicking material into the mould and filling the space inside completely, as temperature went down, it started solidifying. This process took several hours. After that, screws were removed, and the two mould pieces were separated (Figure 4-2 A). A solid tissue-mimicking phantom was then completed (Figure 4-2 B). The volume of the phantom cavity equals to the volume of the semiellipsoidal part of the mould Piece 2. Therefore, by modifying the length of the short- and long-axis of the Piece 2, different phantom cavity volumes were achieved.

The mathematical equation of a standard axis-aligned ellipsoid body in an xyz-Cartesian coordinate system is

$$\frac{x^2}{a}+\frac{y^2}{b}+\frac{z^2}{c}=1$$

where a and b are the equatorial radii (x and y axes) and c is the polar radius (z axis), all of which are fixed positive real numbers determining the shape of the ellipsoid. In this case, a = b < c, which is the prolate spheroid. Three phantoms with different volumes were produced (Figure 4-2 C). The parameters are listed in Table 4-1

Table 4-1 Parameters of three tissue-mimicking phantoms

	Phantom Small	Phantom Medium	Phantom Large
a (mm)	20.0	23.0	35.0
b (mm)	20.0	23.0	35.0
c (mm)	40.0	78.0	80.0
Volume (ml)	33.5	86.4	205.3

The volumes of the Phantom Small and Medium are very typical values for human LVESV and LVEDV. The volume of Phantom Large is much bigger than a normal human LV. The idea of making it was to assess the performance of the new semi-automated algorithms through measuring different volumes. The wall thicknesses of the three phantoms were all 10 mm.

To control the volumes of the phantoms precisely, all mould pieces were cut by an industrial lathe system (XYZ Proturn SLX 1630, Figure 4-3) under CNC (computer numerical control) controller with preset G-code. The material of the moulds is Polymethyl-Methacrylate (PMMA, ignition point 460 ℃, Perspex Ltd.).

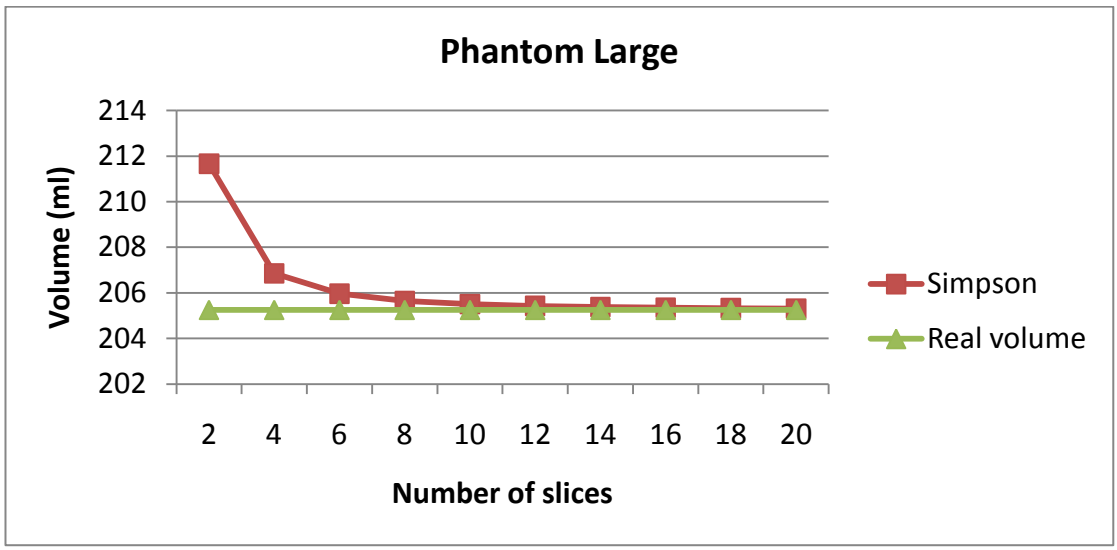
Figure 4-3 XYZ Proturn SLX 1630 lathe system in the workshop of Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne

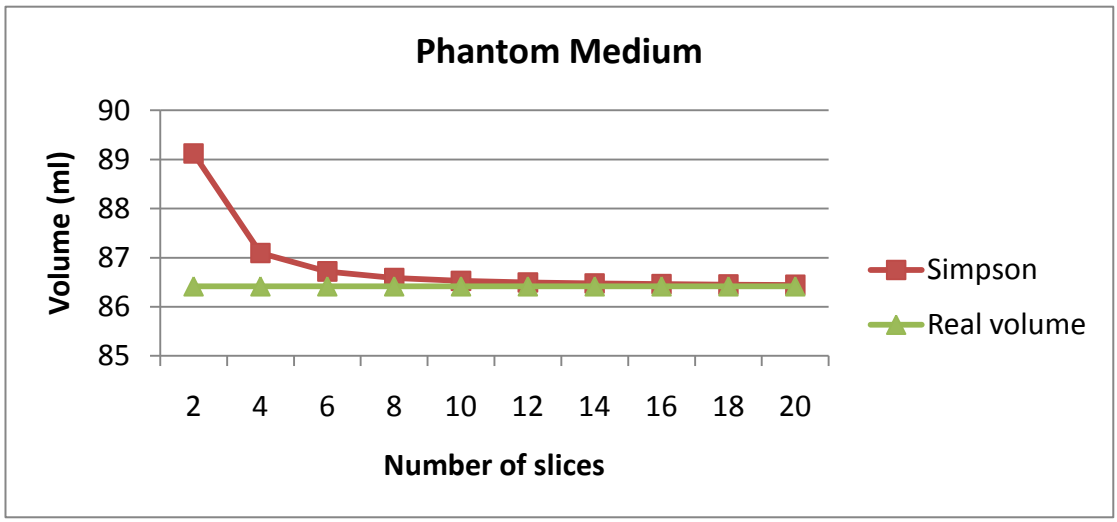
Comparing with previous balloon phantoms, the new tissue-mimicking phantoms do not only offer similar ultrasound property as human myocardium, but also offer standardized geometry and precisely controlled volume.

4.1.3. Imaging acquisition by conventional 2D echocardiography

In this analysis, both 2D and 3D echo images of the phantoms were acquired and compared. To measure a phantom volume, a series of 2D short axis images were acquired by a conventional 2D echocardiography with regular adult transthoracic transducer. Then the summation of discs method (Simpson's rule) was used to calculate the volume. But before doing that, a question needs to be answered first: how many short-axis slices should be acquired to calculate a volume?

It is not difficult to understand that the more slices acquired, the more accurate the calculated volume would be. Figure 4-4 shows a theoretical calculation, which assumes that in an ideal condition, for each phantom volume, the calculated volume by Simpson's rule (red curve) goes closer to the real volume (green line) as the number of





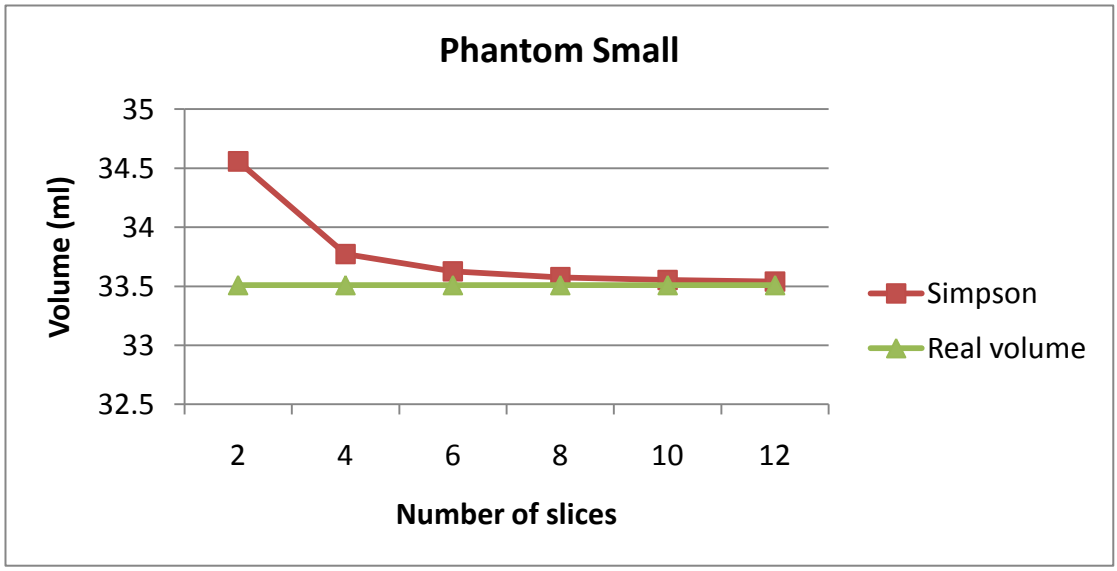


Figure 4-4 Volumes calculated by Simpson's rule in the function of number of slices involved (red curve) for three phantoms respectively. The green line shows the real volume of each phantom.

slices increases. Based on these charts, the slice thickness (which is the distance between the centre of two adjacent slices) was determined as 5mm. This means that in theory, 16 slices can be acquired for the Large and Medium phantom, and 8 slices can be acquired for the Small phantom. Thus the difference between the calculated and real volume is negligible (difference: 0.10ml, 0.04ml and 0.07ml for Phantom Large, Medium and Small respectively).

In order to make sure that the 2D echo transducer was in the right position acquiring the right short axis plane, and every two adjacent planes had the same distance (5 mm), a special imaging acquisition system was used (Figure 4-5).

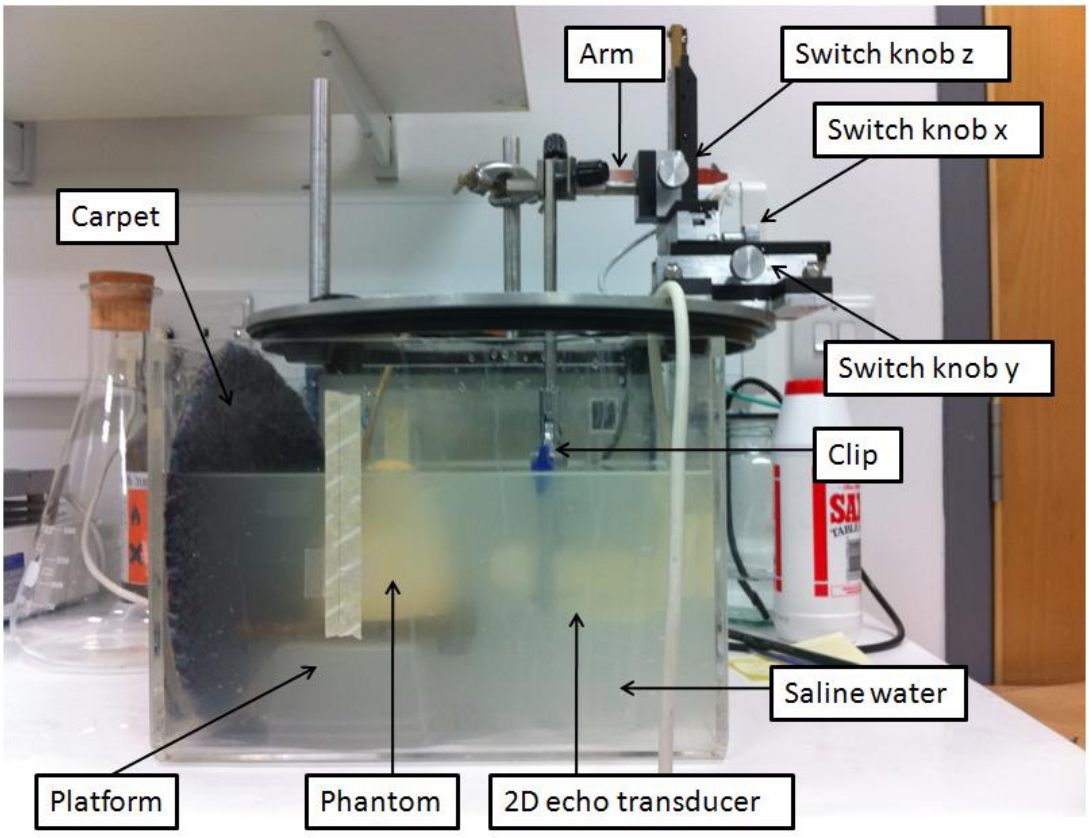


Figure 4-5 2D echo imaging acquisition system

The crucial point for using Simpson's rule in this case was that: the 2D echo transducer should be exactly positioned at 2.5 mm, 7.5 mm, 12.5 mm…and so on, along the long axis of a phantom from its base to apex. To achieve that, a special-made equipment was used to control the vertical movement of the transducer, as shown in Figure 4-5. The 2D echo transducer was held by a clip, and attached to the arm of the equipment with three switch knobs: switch knob x, y and z. By turning them, the arm

can move along x-, y- and z-direction smoothly. For switch knob z, it takes four turns to move the arm 80mm. Therefore, a quarter of one turn can move the transducer 5 mm up or down.

The phantom stood on a metal platform with known height. Thus the start point of imaging acquisition was easily set to 2.5 mm from the phantom base. Then by turning the switch knob z carefully, static 2D echo images of the phantom were captured in every 5 mm movement of the transducer. This procedure was performed for all three phantoms.

Figure 4-5 also shows that both phantom and transducer was submerged inside the saline water instead of pure water in a tank. A study (69) in 2006 done by Y.N.AI-Nassar and his colleagues showed that: the speed of ultrasound in saline water with the salt concentration of 49.5 g/l is 1524 m/s at the temperature of 22 ℃. After a similar test, one of my colleagues David McCulloch found that the ultrasound speed is 1540 m/s in saline water with the concentration of 48 g/l at room temperature, which is exactly the same as for human blood. Because of these tests, the saline water was used to mimic the human blood, so that the volume measurement bias caused by the ultrasound speed difference in different medium can be minimized.

In addition, a carpet was put behind the phantom. The rough surface of the carpet can scatter the ultrasound beam from the transducer. This phenomenon can help to reduce the noise of the 2D echo images of the phantom. Two sets of short axis slices were acquired for each of the phantom. After that, all the images were transferred to a computer for off-line analysis.

4.1.4. Imaging acquisition by Philips 3D echocardiography

The 3D imaging acquisition was performed by the latest Philips iE33 echocardiography system (Figure 4-6). Similar to the 2D imaging acquisition, a tissue-mimicking phantom was submerged inside the saline water. But different from the 2D acquisition, the 3D transducer was hand-held right above the phantom apex. The phantom was then covered by a pyramidal ultrasound field, and a real time 3D image (a dynamic image with a series of continues frames) of the phantom was captured. This was performed at least

twice for each phantom, and all the images were transferred to a computer for off-line analysis.

Figure 4-6 The Philips iE33 echocardiography system

4.2. Imaging analysis

The semi-automated border tracing and volume calculation algorithms were applied to both 2D and 3D echo images to measure the tissue-mimicking phantom volumes. The accuracy as well as the reproducibility of each imaging system was assessed and compared.

4.2.1. Semi-automated measurements on 2D echo images

Six semi-automated algorithms ($G_{1.1}B_1$, $G_{1.1}B_2$, $G_{1.1}B_1I_1$, $G_{1.1}B_2I_1$, $G_{1.1}B_1I_2$ and $G_{1.1}B_2I_2$) were performed on the short axis slices of each phantom. To select the region of interest (ROI), the ROI initialization method 2 was applied (Section 2.1, Figure 2-2): Three slices (phantom apex, middle and base slice, Figure 4-7 A, B and C) were selected for ROI initialization. One point in the phantom cavity and one point in the phantom wall were manually selected for each of the three slices, so that the ROI was established automatically in all slices. Once the ROI was established, the six semi-automated algorithms would start phantom inner wall surface tracing and volume calculation by Simpson's rule.

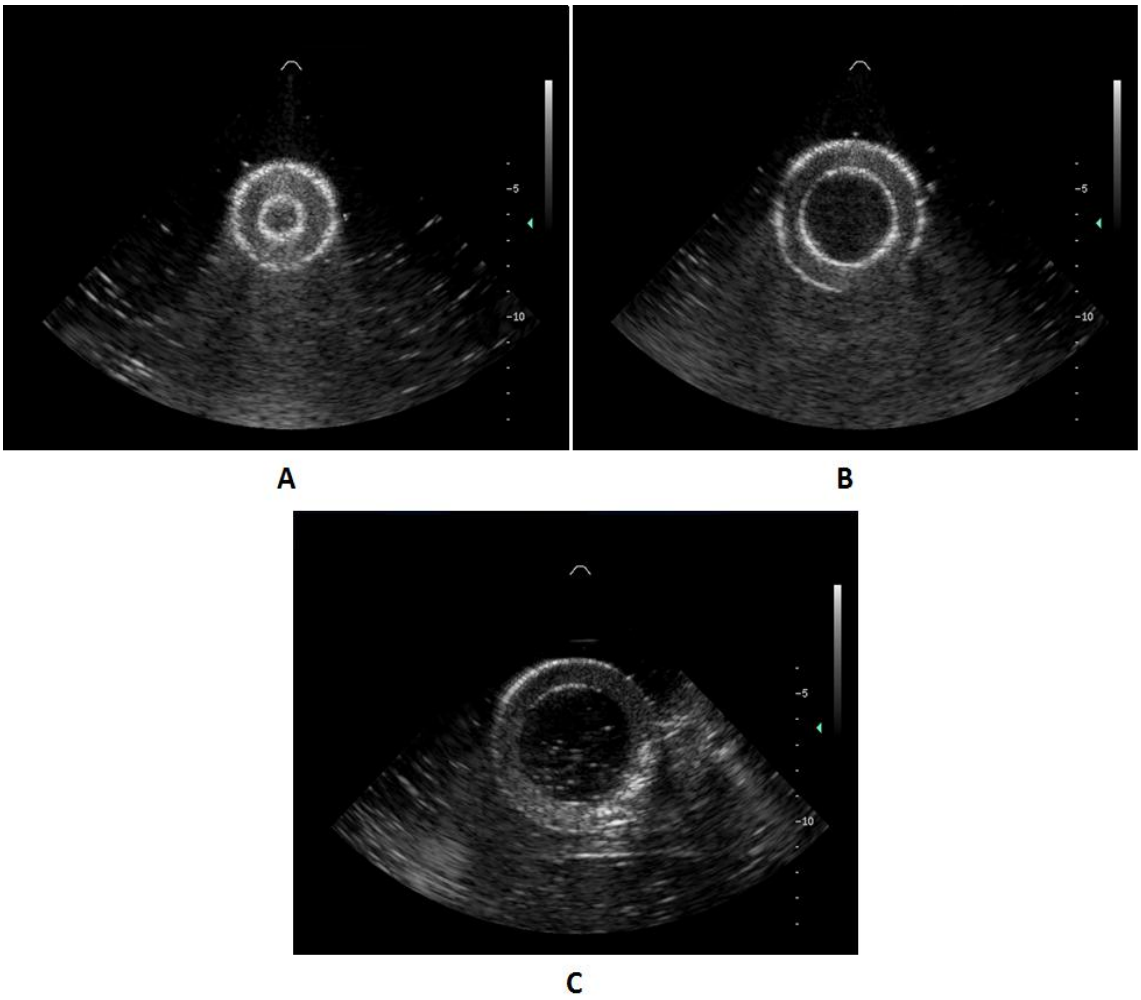


Figure 4-7 Apical (A), middle (B) and basal (C) slice for Phantom Medium, which were acquired by 2D echo. These three slices were selected for ROI initialization.

Since the 2D echo images were discrete slices (5mm gap between each two slices), it was impossible to apply the 3D gradient operator. In this case, only the 5×5 2D gradient operator ($G_{1.1}$) was applied.

4.2.2. Semi-automated measurements on 3D echo images

Similar to the previous analysis in Chapter 3, all the 18 algorithms (Table 3-1) were applied on 3D echo images of tissue-mimicking phantoms. One of the most important findings of the previous analysis was that the 3D operator worked better than the 2D operators. In order to confirm this phenomenon, the FEM was applied on three algorithms this time: $G_{1.1}B_2I_2$, $G_{1.2}B_2I_2$ and $G_2B_2I_2$. Therefore, in total 21 algorithms were used to trace phantom inner surfaces and calculate volumes.

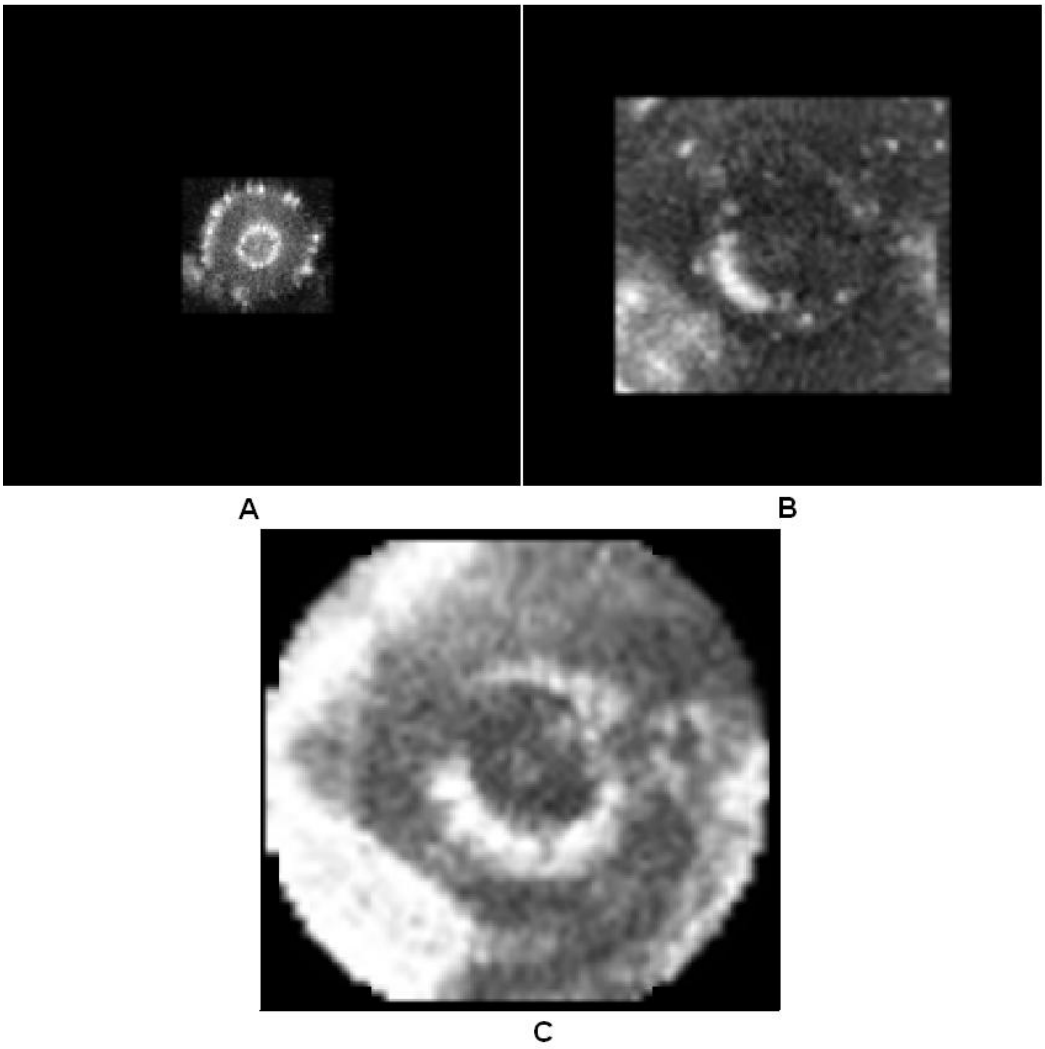


Figure 4-8 Apical (A), middle (B) and basal (C) slice for Phantom Medium, which were acquired by 3D echo. These three slices were selected for ROI initialization.

Similar to the measurements on 2D echo images, the ROI initialization method 2 was also applied on 3D echo images. Figure 4-8 shows an example of the three slices for establishing the ROI for measuring Phantom Medium.

4.2.3. Statistical analysis

For 2D echo imaging analysis, thanks to the precise control of the phantom geometry and the especially designed imaging acquisition system, every 2D image was located to a specific part of a phantom with known volume. Therefore, the measured volume of every slice can be compared with its known volume. The average difference (average bias) and standard deviation (SD) between measured slice volumes and known slice

volumes was calculated for each phantom. The significance of the differences was tested by paired t tests with a two-tailed distribution. $P < 0.05$ was considered significant. The difference between the measured and known phantom volumes was also calculated. Through this way, the accuracy of the semi-automated measurements was assessed.

As mentioned in Section 4.1.3, two sets of 2D short axis slices were captured for every phantom. To assess the reproducibility of the semi-automated algorithms, each algorithm was applied to both sets of images with different manual ROI initializations. The difference between two measurements of one algorithm reflected its variability.

For 3D echo, since the iE33 system captured at least two 3D real time images for each phantom, the image showing the best phantom wall definition was chosen for imaging analysis. The first frame of the real time image was used to apply the semi-automated algorithms.

In a 3D image, one short axis slice volume can be calculated in the slice thickness of one voxel. Similar to the 2D imaging analysis, to assess the accuracy, the average difference and SD between measured and known short axis slice volumes, as well as the difference between the measured and known phantom volumes was calculated for each phantom.

For comparing the slice volumes, paired t tests were applied. $P < 0.05$ was considered significant. Besides that, the ANOVA was employed for all the measurements of slice volumes in order to compare each category of the algorithms ($G_{1.1}$ vs. $G_{1.2}$ vs. G_2, B_1 vs. B_2, no interpolation vs. I_1 vs. I_2) in accuracy.

To assess the reproducibility, for each phantom, the last frame of the chosen image was used to apply all semi-automated algorithms (21 algorithms: 18 algorithms and 3 FEMs) again. Different manual ROI initializations were used for applying each algorithm on two frames. The difference between the measured volumes from the first and the last frame indicated the algorithms variability. To compare the reproducibility of each edge operator, border detection and interpolation algorithm, the ANOVA was also applied.

The FME was applied to $G_{1.1}B_2I_2$, $G_{1.2}B_2I_2$ and $G_2B_2I_2$. The measured slice volumes were compared with known slice volumes in order to assess the accuracy. The difference of measured phantom volumes between the first and last frame was used to assess the reproducibility.

4.3. Results

Theoretically, during 2D echo acquisition, 16 short axis slices should be acquired for Phantom Large and Medium, and eight slices for Phantom Small. However, when the 2D transducer was close to a phantom apex or base, the image quality dropped dramatically. There was no clear appearance of the phantom wall, and thus it was impossible to apply the semi-automated algorithms, even the manual delineations on these images. For this reason, these images were excluded from the analysis. Table 4-2 lists which slices were excluded for each phantom.

Table 4-2 For 2D and 3D echo imaging acquisition of each phantom, slices excluded from imaging analysis in the order of apex to base

	Slices excluded (from apex to base)		
	Phantom Large	**Phantom Medium**	**Phantom Small**
2D echo	1, 2, 14-16	1, 14-16	7, 8
3D echo	All	1, 120-123	1-3, 80-83

For 3D echo images, the imaging quality of Phantom Large was much worse than the other two phantoms. Only the part close to the apex was visible, but the middle and basal part of the phantom was completely lost. Therefore, it was excluded from the analysis. For Phantom Medium and small, the voxel size of the 3D echo images was 0.63mm and 0.48mm respectively. Theoretically, there should be 123 (78 mm/0.63 mm) and 83 (40 mm/0.48 mm) short axis slices captured by 3D echo images of Phantom Medium and Small. However, because of the limitation of the ROI initialization (Section 3.5.2), and the imaging quality issue, several slices close to apex and base were not able to apply the semi-automated algorithms. These slices were also listed in Table 4-2.

4.3.1. Results of 2D echo analysis

For accuracy assessment, Table 4-3 shows the comparison (average difference±SD) between measured slice volumes and corresponding known slice volumes for each phantom in both absolute and relative values. Figure 4-9 shows the curves of slice volumes in the function of slice number given by known volumes and six semi-automated algorithms.

Table 4-3 For 2D echo analysis, the comparison between measured slice volumes and corresponding known slice volumes (average difference±SD) for each phantom in absolute (ml) and relative values (%)

	Measured Slice Volumes - Known Slice Volumes (ml)					
	B_1	B_2	B_1I_1	B_2I_1	B_1I_2	B_2I_2
Phantom Small	-0.5±0.4*	-1.1±0.7*	0.2±0.1*	-0.2±0.1*	0.2±0.1*	-0.2±0.1*
Phantom Medium	-0.2±0.3	-1.4±0.8*	0.0±0.2	-0.4±0.3*	0.1±0.2	-0.4±0.2*
Phantom Large	-1.2±1.2*	-6.2±1.1*	0.1±1.1	-0.6±0.9	0.4±1.2	-0.4±1.1
	(Measured Slice - Known Slice Volumes)/Know Slice Volumes (%)					
	B_1	B_2	B_1I_1	B_2I_1	B_1I_2	B_2I_2
Phantom Small	-8.1±14.4	-29.9±7.1*	8.4±8.9	-4.6±1.8*	9.8±8.9*	-3.5±1.0*
Phantom Medium	-0.1±8.5	-24.9±6.3*	3.7±7.8	-7.7±2.8*	4.1±7.6	-5.8±3.6*
Phantom Large	-13.4±15.1*	-51.5±16.2*	-1.9±12.2	-7.4±12.3	0.1±12.2	-5.9±13.7

Note: '*' means significant difference.

Table 4-4 For 2D echo analysis, the comparison between measured phantom volumes and known phantom volumes

	Measured Phantom Volume - Known Phantom Volume (ml)					
	$G_{1.1}B_1$	$G_{1.1}B_2$	$G_{1.1}B_1I_1$	$G_{1.1}B_2I_1$	$G_{1.1}B_1I_2$	$G_{1.1}B_2I_2$
Phantom Small	-15.1	-19.2	-11.3	-13.2	-11.0	-13.1
Phantom Medium	-26.7	-41.4	-24.2	-30.1	-23.9	-29.1
Phantom Large	-75.4	-129.7	-60.2	-67.9	-57.4	-65.6
Average	-39.1	-63.4	-31.9	-37.1	-30.8	-35.9

Note: The missed volumes due to excluded slices were 12.3ml, 24.7ml and 61.7ml for Phantom Small, Medium and Large

From Table 4-3 and Figure 4-9, it is obvious that for acquired short axis slices, all the semi-automated algorithms showed small differences with known slice volumes except the algorithm $G_{1.1}B_2$, which showed the biggest bias for all three phantoms (Table 4-3 column 2). The interpolation algorithms significantly improved the measurements, especially for B_2. In most of the cases, the biases of algorithms with B_1 (Table 4-3,

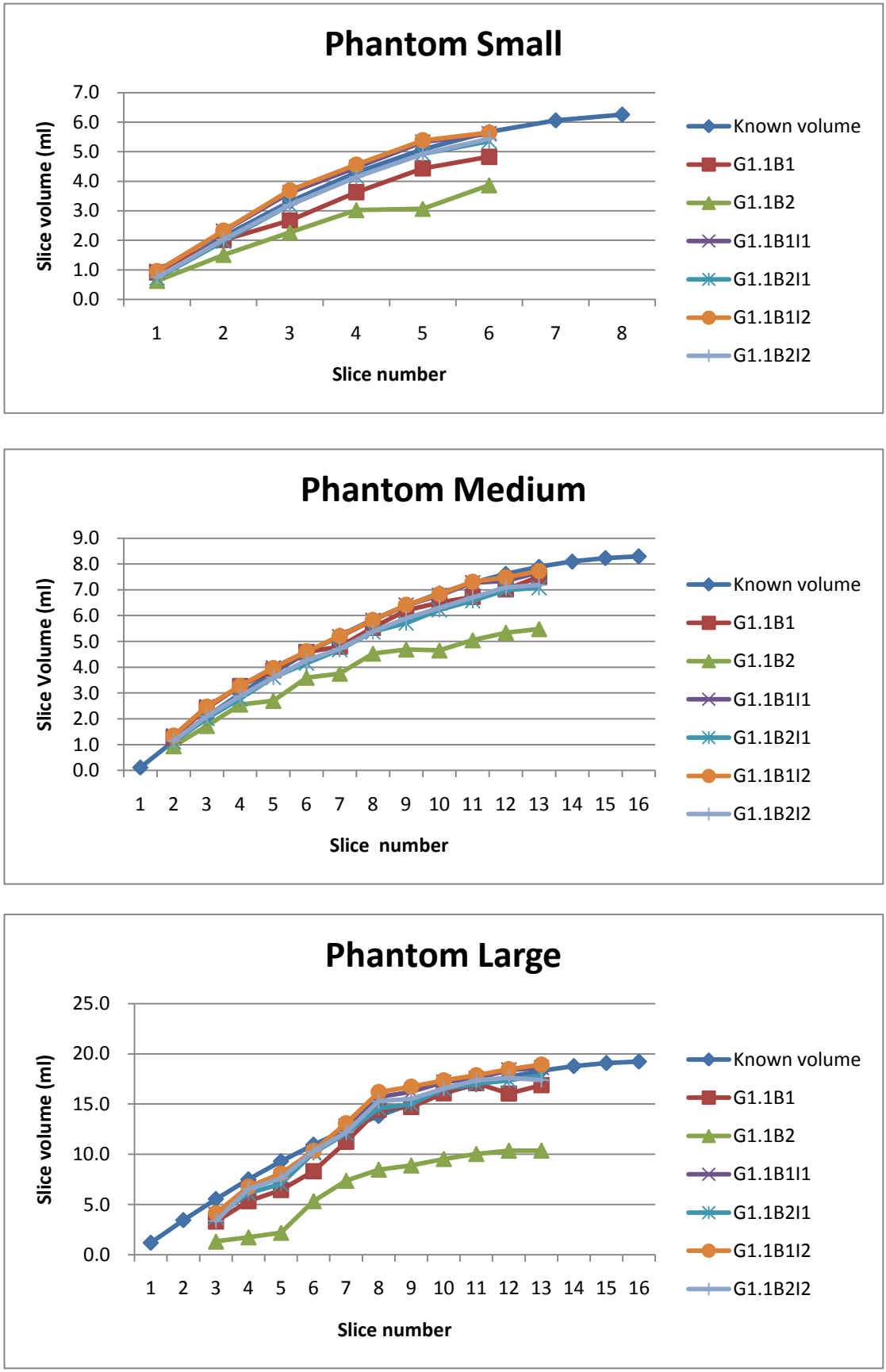


Figure 4-9 Curves of slice volumes in the function of slice number from apex to base of each tissue-mimicking phantom

column 1, 3 and 5) were smaller than or the same as the biases of algorithms with B_2 (Table 4-3, column 2, 4 and 6), except the measurements of Phantom Small in relative value.

Table 4-4 lists the differences between measured phantom volumes and known phantom volumes. The biggest average bias was -63.4 ml, given by $G_{1.1}B_2$, and the smallest average bias was -30.8 ml, given by $G_{1.1}B_1I_2$. It should be noted that the reason of such big underestimation of the phantom volumes was the missing of several short axis slices caused by unacceptable image quality. As shown in Table 4-2, for each phantom, two to five (Phantom Small to Phantom Large) short axis images were not captured for measurements. The missed volumes due to these excluded slices were 12.3ml, 24.7ml and 61.7ml for Phantom Small, Medium and Large respectively. If these volumes were added into the comparison, the biases of all the semi-automated measurements would be much smaller.

Table 4-3 and Figure 4-9 also proves that for acquired slices (slices could be measured by the algorithms), the biases between measured slice volumes and known slice volumes were very small (smallest bias: 0.1%). Therefore, for measuring phantom volumes, 2D echocardiography showed big degree of systematic bias because of not able to acquire enough short axis slices.

Table 4-5 Difference between the first and second volume measurements by six semi-automated algorithms

	Second Measurement - First Measurement (ml)					
	$G_{1.1}B_1$	$G_{1.1}B_2$	$G_{1.1}B_1I_1$	$G_{1.1}B_2I_1$	$G_{1.1}B_1I_2$	$G_{1.1}B_2I_2$
Phantom Small	-0.2	-0.3	-1.0	-1.0	-0.9	-1.1
Phantom Medium	1.3	0.6	1.8	1.5	1.8	2.0
Phantom Large	0.8	-1.2	-3.0	0.4	-2.0	1.0
Average	0.6	-0.3	-0.8	0.3	-0.3	0.6

For reproducibility assessment, Table 4-5 shows the difference between the first and second volume measurements given by six algorithms. The differences were very small in all cases. The maximal difference was -3.0 ml given by $G_{1.1}B_1I_1$ for measuring Phantom Large. This test proved that the semi-automated algorithms were very reproducible in quantitative volume measurement.

4.3.2. Results of 3D echo analysis

For accuracy assessment, Table 4-6 shows the performance of the 21 algorithms (including three FEMs) by comparing the differences of measured slice volumes and known slice volumes of each phantom. Because the thickness of every slice was just one voxel, whose thickness was 0.48mm and 0.63mm for Phantom Small and Medium, the volume of one slice was very small. Take the Phantom Small as an example. The volume of Phantom Small was 33.5ml (Table 4-1), and there were 83 slices, thus the average volume of each slice was about 0.4ml (33.5ml/83). Therefore, the differences between measured slice volumes and known slice volumes were very small. It should be noted that the unit of Table 4-6 is 0.1ml instead of 1ml.

Table 4-6 For 3D echo analysis, the comparison between measured slice volumes and corresponding known slice volumes (average difference±SD) of each phantom with 21 algorithms in both absolute (10^{-1} ml) and relative values (%)

	Measured Slice Volumes - Known Slice Volumes (10^{-1} ml)						
Phantom	**$G_{1.1}B_1$**	**$G_{1.1}B_2$**	**$G_{1.1}B_1I_1$**	**$G_{1.1}B_2I_1$**	**$G_{1.1}B_1I_2$**	**$G_{1.1}B_2I_2$**	**$G_{1.1}B_2I_2$+FEM**
Small	-1.2±1.5*	-1.0±1.5*	-0.3±1.7	-0.0±1.6	-0.1±1.7	-0.0±1.6	-0.0±1.6
Medium	-2.3±2.3*	-2.9±2.2*	-0.5±2.4*	0.3±2.6	-0.5±2.3*	-0.0±2.6	-0.0±2.3
	$G_{1.2}B_1$	**$G_{1.2}B_2$**	**$G_{1.2}B_1I_1$**	**$G_{1.2}B_2I_1$**	**$G_{1.2}B_1I_2$**	**$G_{1.2}B_2I_2$**	**$G_{1.2}B_2I_2$+FEM**
Small	-1.6±1.5*	-1.1±1.5*	-0.4±1.6	-0.1±1.6	-0.3±1.6	-0.1±1.6	-0.1±1.6
Medium	-2.8±2.2*	-3.1±2.1*	-0.6±2.3*	0.1±2.5	-0.7±2.3*	-0.2±2.6	-0.2±2.3
	G_2B_1	**G_2B_2**	**$G_2B_1I_1$**	**$G_2B_2I_1$**	**$G_2B_1I_2$**	**$G_2B_2I_2$**	**$G_2B_2I_2$+FEM**
Small	-1.4±1.5*	-1.1±1.4*	-0.2±1.6*	-0.0±1.6	-0.0±1.8	-0.0±1.6	-0.0±1.6
Medium	-2.8±2.3*	-3.0±2.2*	-0.6±2.4*	0.3±2.6	-0.5±2.4*	0.1±2.6	-0.0±2.3
	(Measured Slice - Known Slice Volumes)/Known Slice Volumes (%)						
Phantom	**$G_{1.1}B_1$**	**$G_{1.1}B_2$**	**$G_{1.1}B_1I_1$**	**$G_{1.1}B_2I_1$**	**$G_{1.1}B_1I_2$**	**$G_{1.1}B_2I_2$**	**$G_{1.1}B_2I_2$+FEM**
Small	-27.3±26.1*	-39.6±23.3*	-2.4±28.8	2.7±32.5	-4.2±30.3	3.6±33.5	2.0±29.5
Medium	-32.3±26.7*	-39.0±21.2*	-8.3±27.4*	1.4±28.5	-7.6±25.5*	-2.4±28.5	-5.1±28.4
	$G_{1.2}B_1$	**$G_{1.2}B_2$**	**$G_{1.2}B_1I_1$**	**$G_{1.2}B_2I_1$**	**$G_{1.2}B_1I_2$**	**$G_{1.2}B_2I_2$**	**$G_{1.2}B_2I_2$+FEM**
Small	-43.2±27.4*	-27.2±29.1*	-8.8±35.9*	-2.5±35.7	-7.1±34.6	-3.0±35.4	-3.8±34.6
Medium	-40.4±23.7*	-44.2±20.0*	-12.9±28.0*	-4.2±29.9	-9.7±52.1*	-3.6±53.9	-9.8±27.4*
	G_2B_1	**G_2B_2**	**$G_2B_1I_1$**	**$G_2B_2I_1$**	**$G_2B_1I_2$**	**$G_2B_2I_2$**	**$G_2B_2I_2$+FEM**
Small	-38.2±29.6*	-27.5±28.7*	-7.3±40.4	0.6±37.0	4.9±74.8	0.8±36.8	0.4±36.1
Medium	-34.4±27.2*	-41.8±20.6*	-6.7±28.0*	4.9±28.0	-5.9±25.5*	0.4±28.6	-0.7±24.7

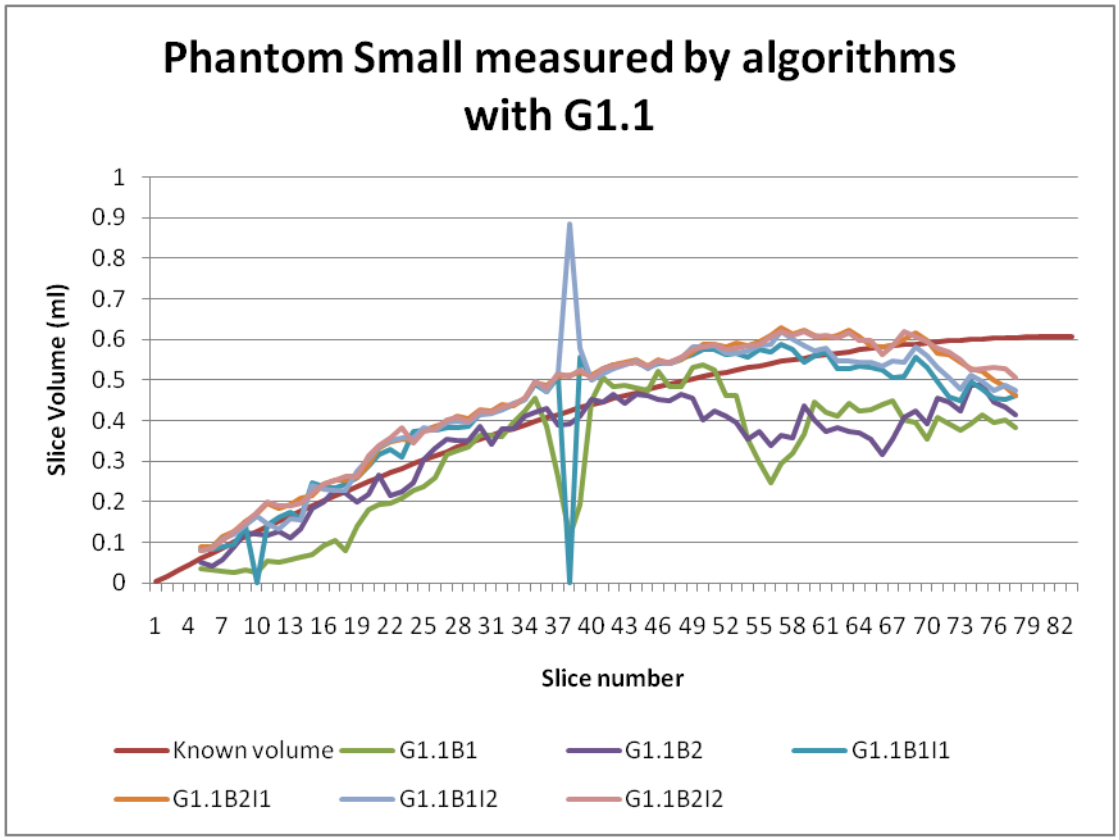
Note: '*' means significant difference. The unit of this table is 10^{-1} ml.

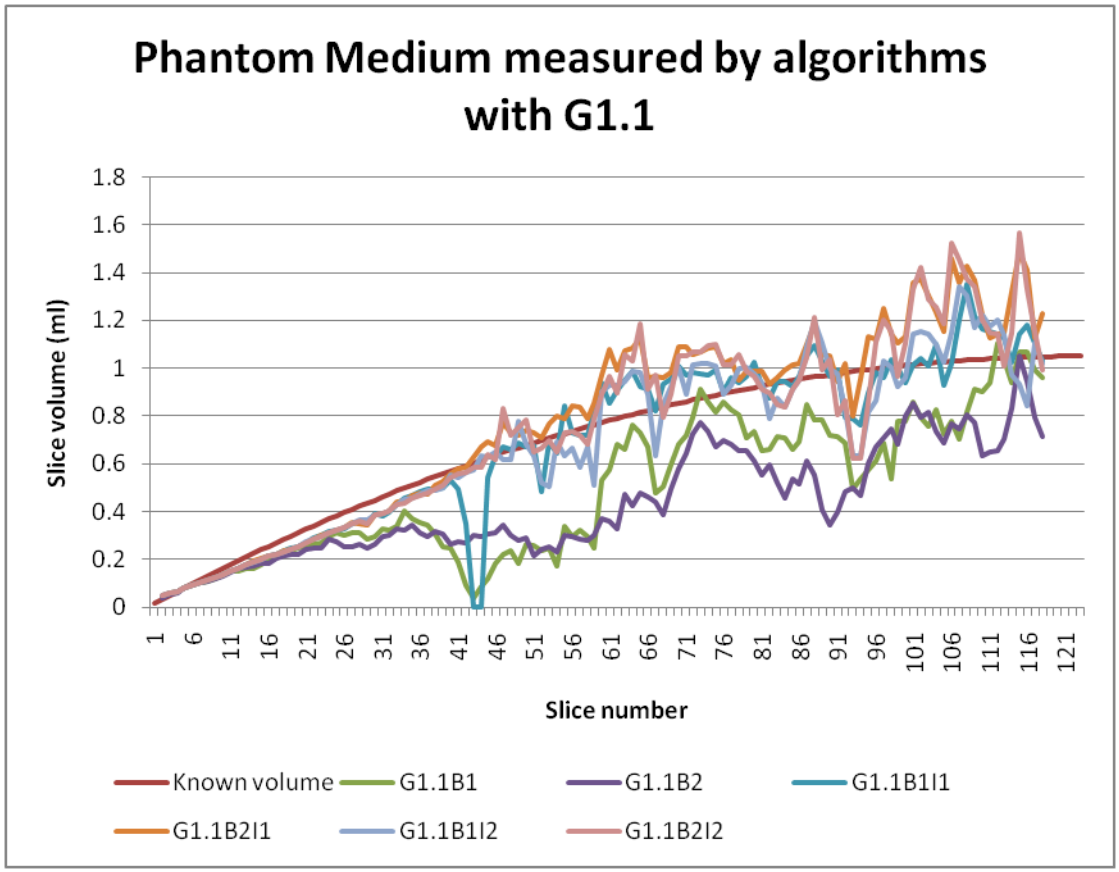
Similar to the 2D echo analysis, the average biases of all measurements were small. The biggest bias was -0.31±0.21 ml, given by $G_{1.2}B_2$ for measuring Phantom Medium. The interpolation algorithms (I_1/I_2) improved the accuracy as well. All the measurements without interpolation (Table 4-6 column 1 and 2) showed significant difference with known volumes. However, Table 4-3 and Table 4-6 indicate that there were two differences between the 2D and 3D echo analyses.

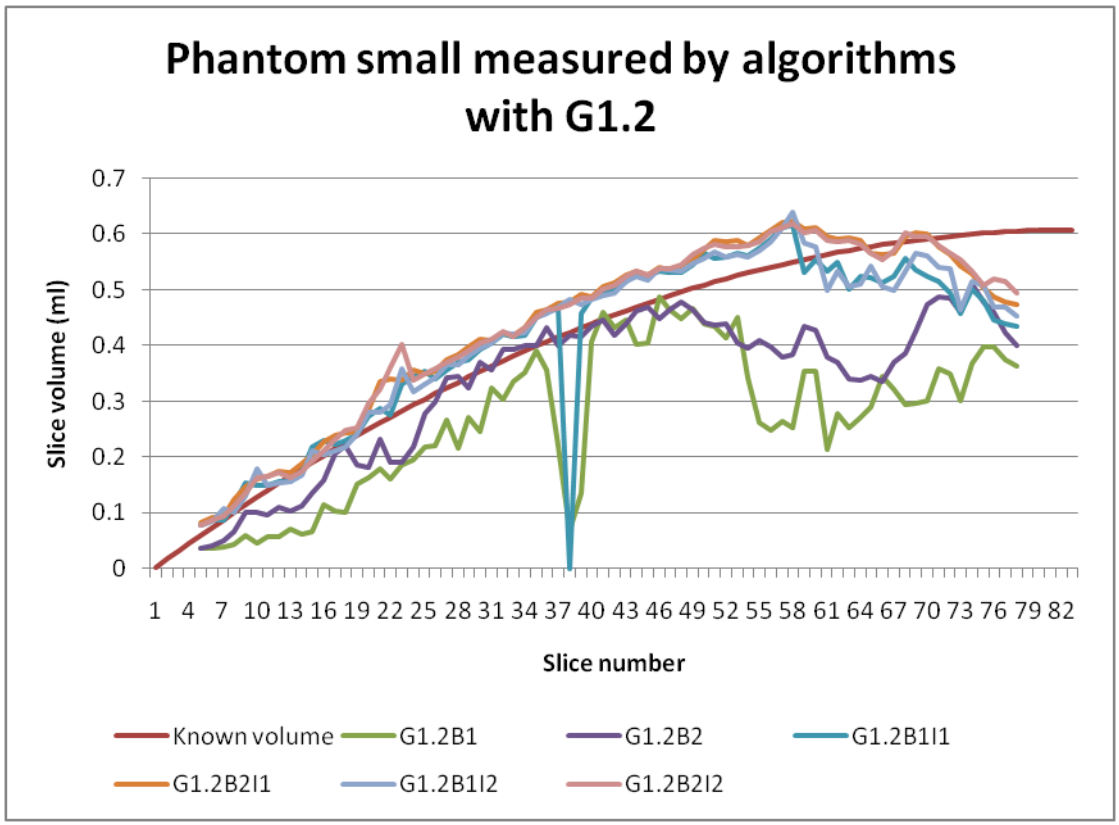
Firstly, different from the 2D echo analysis, with the help of interpolation algorithms, B_2 demonstrated better results than B_1 did in 3D images. Table 4-6 indicates that all the algorithms with B_2 plus interpolations (Table 4-6 column 4 and 6) showed no significant difference with known volumes, but B_1 plus interpolations (Table 4-6 column 3 and 5) showed significant difference in many cases. Secondly, if the comparison was in relative values, although the average biases of all measurements were small in both analyses, the algorithms showed much bigger standard deviation (SD) in 3D echo analysis. The biggest SD in Table 4-3 is 16.2% from B2 measuring Phantom Large, which is still smaller than all the SD in Table 4-6.

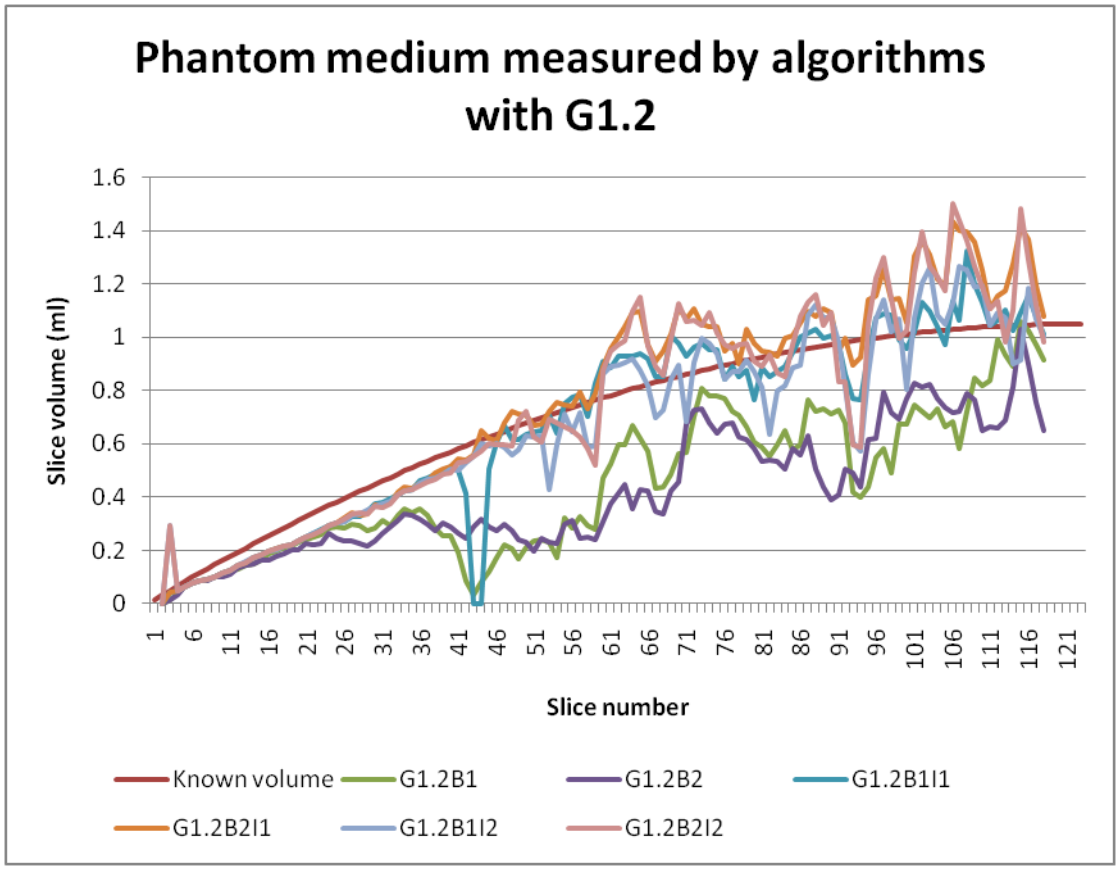
After applying the FEM, the results showed equal or less biases when comparing in absolute values, but not in relative values (Table 4-6 column 6 and column 7). However, comparing with the original algorithms, the FEM reduced the SD for all the measurements in both absolute and relative values.

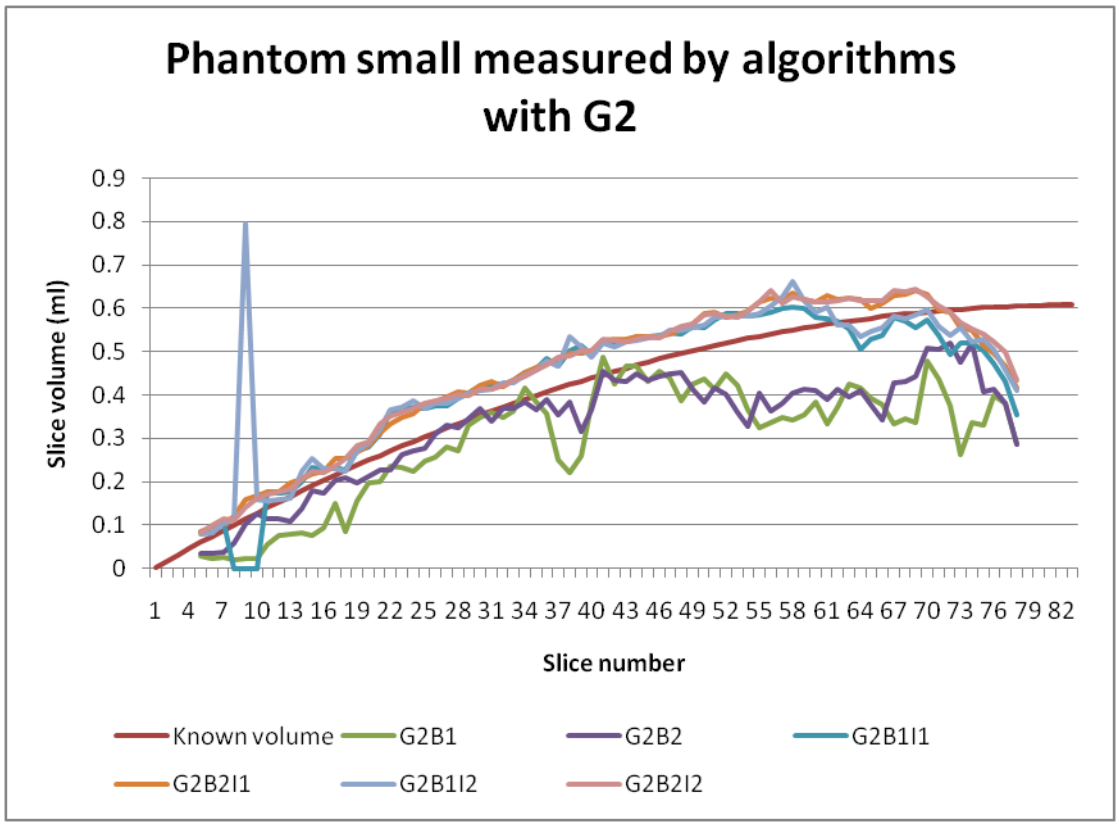
Figure 4-10 shows the curves of slice volumes as a function of slice number given by known volumes and 18 semi-automated algorithms (FEM not included). These plots confirm that the curves of algorithms with interpolations were following the reference curves more closely. However, Figure 4-10 also shows that the border tracing was not always successful. Even with the help of interpolations, there were still some extreme biases appearing on certain slices. For example, $G_{1.1}B_1I_1$ and $G_{1.1}B_1I_2$ were failed in measuring the slice 38 of Phantom Small (Figure 4-10 the first plot). This was one of the reasons that caused big standard deviations while comparing measured slice volumes with known slice volumes in relative values (Table 4-6).


Phantom Small measured by algorithms with G1.1
Slice Volume (ml)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82
Slice number
Known volume
G1.1B1
G1.1B2
G1.1B1I1
G1.1B2I1
G1.1B1I2
G1.1B2I2


Phantom Medium measured by algorithms with G1.1
Slice volume (ml)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 76 81 86 91 96 101 106 111 116 121
Slice number
Known volume
G1.1B1
G1.1B2
G1.1B1I1
G1.1B2I1
G1.1B1I2
G1.1B2I2


Phantom small measured by algorithms with G1.2
Slice volume (ml)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82
Slice number
Known volume
G1.2B1
G1.2B2
G1.2B1I1
G1.2B2I1
G1.2B1I2
G1.2B2I2


Phantom medium measured by algorithms with G1.2
Slice volume (ml)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 76 81 86 91 96 101 106 111 116 121
Slice number
Known volume
G1.2B1
G1.2B2
G1.2B1I1
G1.2B2I1
G1.2B1I2
G1.2B2I2



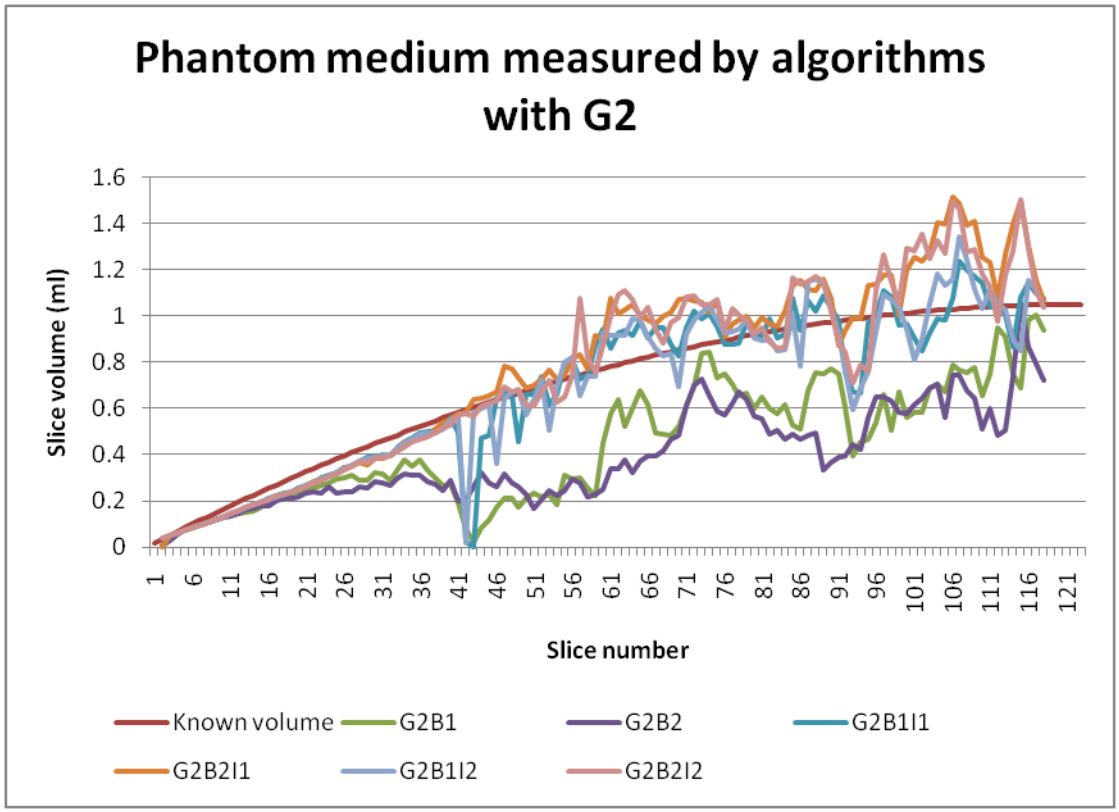


Figure 4-10 Curves of slice volumes in the function of slice number from phantom apex to base for 18 algorithms (FEM not included)

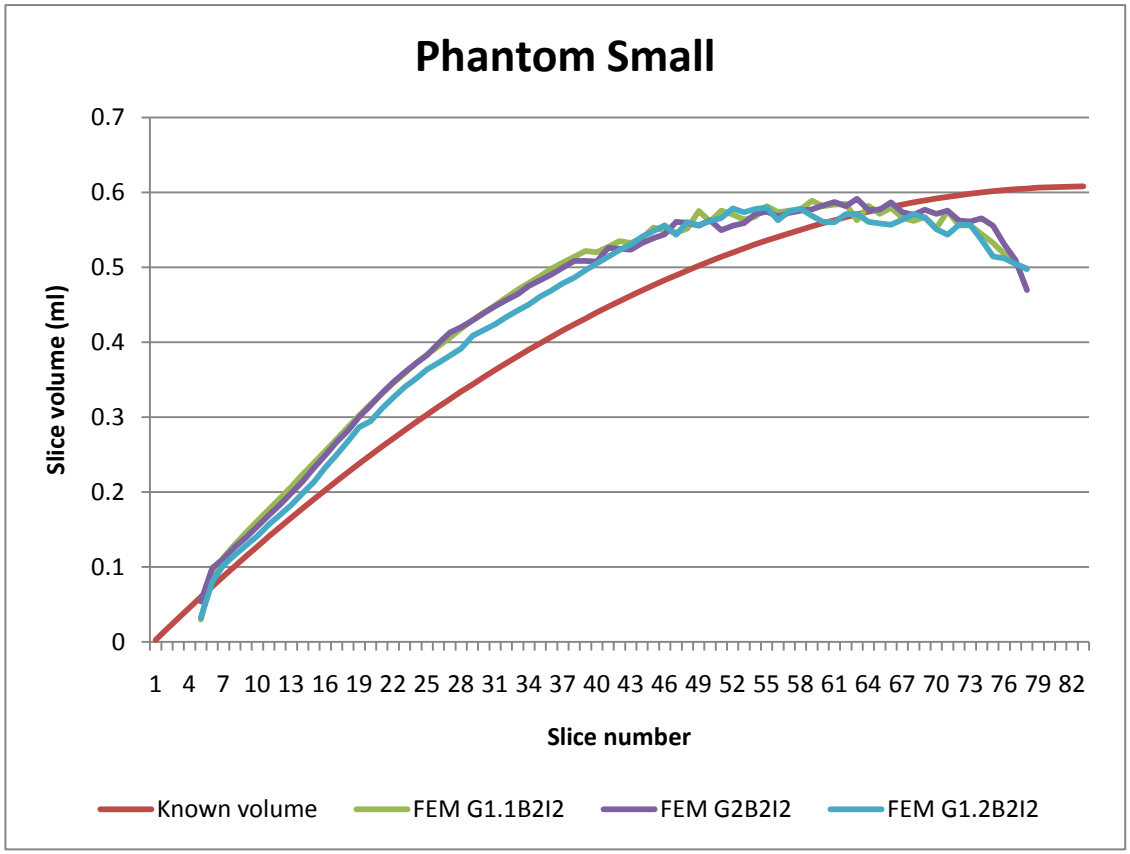
Table 4-7 lists the differences between measured phantom volumes and known phantom volumes given by 21 algorithms. This table clearly shows again that the interpolation algorithms improved the measurements significantly. Same as the 2D echo analysis, there was also a certain degree of systematic bias because of excluding slices from the measurements in the 3D echo analysis. However, since the 3D echo images offered more information of the areas close to phantom apex and base, the missed slice volumes were much smaller this time. They were only 3.1ml and 5.3ml for Phantom Small and Medium. Same as the Table 4-4, Table 4-7 did not take this systematic bias into account. If these volumes were added into the comparison, the biases of all measurements in Table 4-7 would be averagely 4ml bigger.

Table 4-7 For 3D echo analysis, the comparison between measured phantom volumes and known phantom volumes of 21 algorithms

	Measured Phantom Volumes - Known Phantom Volumes (ml)						
Phantom	**$G_{1.1}B_1$**	**$G_{1.1}B_2$**	**$G_{1.1}B_1I_1$**	**$G_{1.1}B_2I_1$**	**$G_{1.1}B_1I_2$**	**$G_{1.1}B_2I_2$**	**$G_{1.1}B_2I_2$+FEM**
Small	-10.3	-8.5	-2.7	-0.2	-1.0	-0.1	-0.4
Medium	-28.8	-36.0	-5.9	4.1	-5.8	0.1	-1.0
Average	-19.6	-22.3	-4.3	1.9	-3.4	0.0	-0.7
	$G_{1.2}B_1$	**$G_{1.2}B_2$**	**$G_{1.2}B_1I_1$**	**$G_{1.2}B_2I_1$**	**$G_{1.2}B_1I_2$**	**$G_{1.2}B_2I_2$**	**$G_{1.2}B_2I_2$+FEM**
Small	-13.7	-8.7	-3.3	-1.1	-2.7	-1.2	-1.3
Medium	-34.3	-37.9	-7.6	1.6	-8.7	-1.9	-2.9
Average	-24.0	-23.3	-5.5	0.2	-5.7	-1.5	-2.1
	G_2B_1	**G_2B_2**	**$G_2B_1I_1$**	**$G_2B_2I_1$**	**$G_2B_1I_2$**	**$G_2B_2I_2$**	**$G_2B_2I_2$+FEM**
Small	-11.5	-9.0	-2.0	-0.2	-0.4	-0.1	-0.4
Medium	-34.0	-37.1	-6.9	4.1	-6.9	1.1	-0.1
Average	-22.8	-23.0	-4.5	1.9	-3.6	0.5	-0.3

Note: The missed volumes due to excluded slices were 3.1ml and 5.3ml for Phantom Small and Medium.

Comparing with the original algorithms (Table 4-7 column 6), the FEM made all biases more negative by 1.2ml at most (Table 4-7 column 7). Figure 4-11 shows the curves of slice volume vs. slice number for the three FEM algorithms. Comparing with Figure 4-10, these curves were smoother, and there was no failed measurement on any slice of both phantoms.



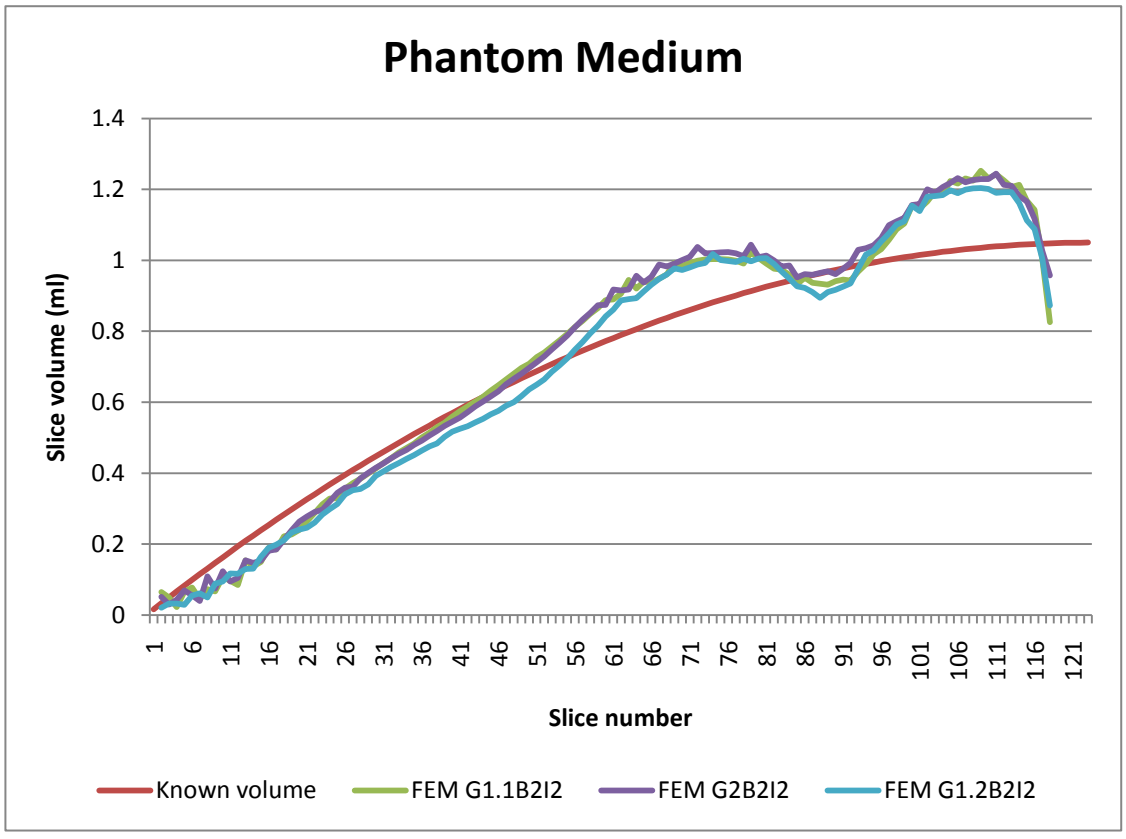


Figure 4-11 Curves of slice volumes in the function of slice number from phantom apex to base for three algorithms ($G_{1.1}B_2I_2$, $G_{1.2}B_2I_2$ and $G_2B_2I_2$) applied FEM

After applying ANOVA on the measured phantom slice volumes from the 18 semi-automated algorithms, the performance of each edge operator, border detection and border interpolation algorithm was assessed.

For three edge operators, the results showed that there was no significant difference between $G_{1.1}$ and G_2 (P=0.23; mean bias with known slice volumes: $G_{1.1}$: -0.07 ml and G_2: -0.07 ml). However, there was a significant difference between $G_{1.1}$ and $G_{1.2}$ (P<0.05), as well as G_2 and $G_{1.2}$ (P<0.05; mean bias with known slice volumes: $G_{1.2}$: -0.09 ml). For two border detection algorithms, the results showed significant difference (P<0.05; mean bias with known slice volumes: B_1: -0.10ml and B_2: -0.07 ml). For interpolation algorithms, although no significant difference was found between I_1 and I_2 (P=0.52; mean bias with known slice volumes: I_1: -0.01 ml and I_2: -0.02 ml), there was a significant difference between with and without interpolations (P<0.05; mean bias with known slice volumes: no interpolation: 0.22 ml). This result was consistent with the result of balloon phantom analysis in Section 3.3.1.

Table 4-8 Difference between the volume measurements of the first and the last frame of Phantom Medium and Phantom Small, given by the 21 semi-automated algorithms

	Measurement of last frame - Measurement of first frame (ml)						
Phantom	**$G_{1.1}B_1$**	**$G_{1.1}B_2$**	**$G_{1.1}B_1I_1$**	**$G_{1.1}B_2I_1$**	**$G_{1.1}B_1I_2$**	**$G_{1.1}B_2I_2$**	**$G_{1.1}B_2I_2$+FEM**
Small	1.0	0.5	1.1	1.3	4.2	1.1	1.1
Medium	13.7	9.1	11.4	12.7	8.8	6.6	6.5
Average	7.3	4.8	6.3	7.0	6.5	3.9	3.8
	$G_{1.2}B_1$	**$G_{1.2}B_2$**	**$G_{1.2}B_1I_1$**	**$G_{1.2}B_2I_1$**	**$G_{1.2}B_1I_2$**	**$G_{1.2}B_2I_2$**	**$G_{1.2}B_2I_2$+FEM**
Small	0.9	0.2	0.8	0.7	1.4	0.8	0.4
Medium	7.8	3.8	5.6	5.9	3.5	2.9	2.8
Average	4.4	2.0	3.2	3.3	2.4	1.9	1.6
	G_2B_1	**G_2B_2**	**$G_2B_1I_1$**	**$G_2B_2I_1$**	**$G_2B_1I_2$**	**$G_2B_2I_2$**	**$G_2B_2I_2$+FEM**
Small	-0.8	-0.2	-0.8	0.2	-1.0	-0.2	-0.1
Medium	7.1	4.7	6.1	4.7	3.6	3.7	3.4
Average	3.1	2.3	2.6	2.5	1.3	1.7	1.6

For the reproducibility assessment, the difference of the volume measurements between the first and last frame of phantom Medium and Small, given by the 21 semi-automated algorithms, were listed in Table 4-8.

After applying the ANOVA on the differences of the two frame volume

measurements given by the 18 algorithms (FEM not included), significant differences were detected between the edge operator $G_{1.1}$ and $G_{1.2}$ ($P<0.05$; Mean of the frame differences: $G_{1.1}$=5.0 ml; $G_{1.2}$=2.9 ml), as well as $G_{1.1}$ and G_2 ($P<0.05$; Mean of the frame differences: G_2=3.0 ml). But there was no significant difference between $G_{1.2}$ and G_2 (P=0.85). There was neither significant difference between border detection algorithms (P=0.32; Mean of the frame differences: B_1=3.9 ml; B_2=3.3 ml), nor significant difference for using or not using interpolation algorithms: P=0.86 for no interpolation and using I_1; P=0.17 for no interpolation and using I_2; P=0.12 for I_1 and I_2. Mean of the frame differences: no interpolation=3.9 ml; I_1=4.1 ml; I_2=2.7 ml.

Comparing with the original algorithms (Table 4-8 column 6), after applying the FEM, the differences of two frame volume measurements were the same or slightly smaller (Table 4-8 column 7). There was no significant improvement of the reproducibility by using the FEM.

4.4. Discussion

In the previous 3D echo imaging analysis of balloon phantoms in Chapter 3, the accuracy and reproducibility of the newly developed 18 semi-automated algorithms and the FEM was assessed and compared. However, there were some inherent drawbacks of the previous analysis, such as the issue of ultrasound speed differences in different materials and the non-precisely control of phantom geometry and volume.

To overcome these drawbacks, in this chapter, great efforts were invested into the design and production of the new tissue-mimicking phantoms, and the saline water concentration control, so that the new phantoms and saline has the same ultrasound speed as the human soft tissue and blood. Besides that, a special imaging acquisition system was used to acquire short axis 2D images by conventional echocardiography. With the help of this system, the differences of performing semi-automated quantitative volume measurements on 2D and 3D images can be compared.

By applying the semi-automated algorithms on 3D images of the new phantoms, the conclusions of the previous analysis were verified. More evidence was obtained to

support the findings of Chapter 3.

This section discusses the problems and the findings of the tissue-mimicking phantom analysis, the strengths and weaknesses of the 2D and 3D echo images, and the meanings of this investigation for clinical applications.

4.4.1. What influenced the image quality?

After the 2D echo imaging acquisition process, some images were excluded from the analysis (Table 4-2) because of the bad quality. Those images were close to phantom apex or base. For 3D echo images, the same problem occurred (see Table 4-2). Figure 4-7 and Figure 4-8 shows that the quality of the basal slices chosen for ROI initialization was worse than the quality of the other two slices. What caused this phenomenon?

The biggest interference of the 2D echo image quality was the strong ultrasound reflection from the metal platform and the tank wall. When the 2D echo transducer was close to the phantom base or apex, the ultrasound signal reflected from the tank wall or metal platform caused strong artefact. This resulted in losing the phantom wall boundaries in echo images. To reduce the ultrasound reflection from the tank wall, a piece of carpet was placed behind the phantom during the imaging acquisition, as shown in Figure 4-5. Without this carpet, even for the middle part of the phantom, the image quality would be much worse. However, this did not reduce the interference on the phantom apical and basal images.

Different from 2D echo acquisition, 3D echo transducer was held vertically above the phantom apex. The same carpet was placed under the phantom. However, the ultrasound signal reflected from the carpet and the tank wall still influenced the image quality of the phantom around its basal area, which caused the appearance of noise and artefact. Figure 4-8 C shows that phantom basal wall boundary was partially lost and totally distorted. The short axis contour should be a perfect circle, which is true in Figure 4-8 A and B. But in C, the contour turns to be an ellipse.

It should be noted that the reason of excluding slices around the phantom apical area was not because of the image quality for 3D echo images. In fact, the 3D echo images

provided excellent image quality around the phantom apical area, which the 2D echo was not capable of. The limitation of the ROI initialization (Section 3.5.2) caused a small part of the apical area to be excluded from the volume calculation.

Another interference of the image quality was the micro air bubbles. The air bubbles showed very strong ultrasound reflection, and were captured as very bright small dots in the echo images (Figure 4-12 A). It increased the image noise, and might cause some wrong detection of the phantom wall boundaries. But since it did not introduce significant artefact, and its influence was minimized by the edge operators, interpolation algorithms and FEM, its appearance was not considered for the exclusion criteria.

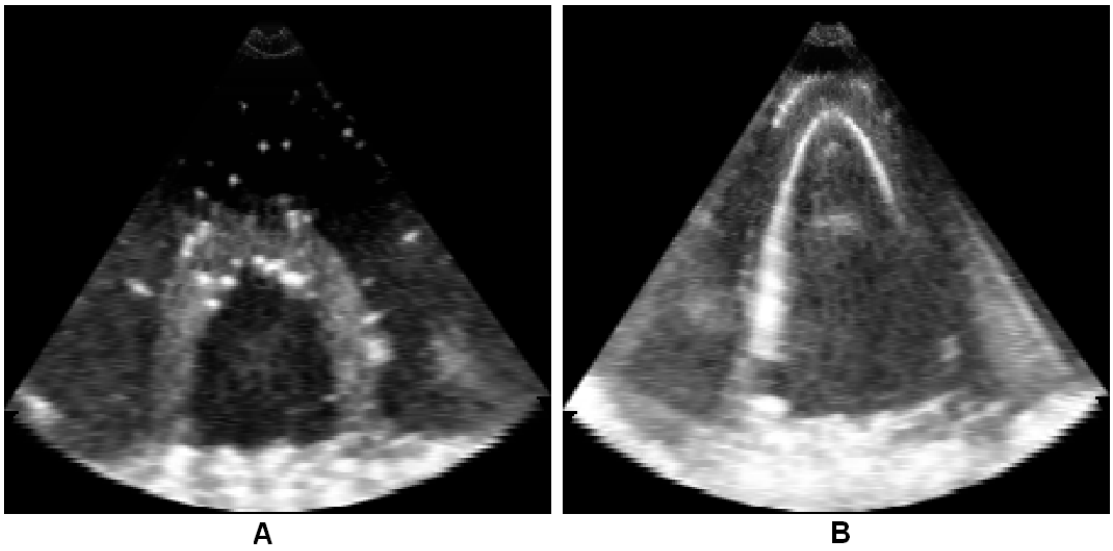


Figure 4-12 Long axis views of the Phantom Small (A) and Medium (B), extracted from the 3D echo images of both phantoms

The 3D echo image of the Phantom Large was excluded from the measurement. Comparing with the 3D images of the other two phantoms, the three acquired 3D images of the Phantom Large all showed unexpected bad quality. The reasons of that need to be further investigated. One of the possible explanations was that the 3D echo transducer provides better image quality if the scanning field is smaller.

The 3D echo transducer creates an ultrasound scanning field of pyramidal shape. If the subject is small, the 3D scanning field is small. Therefore, the concentration of the ultrasound signal in the field is higher. The signal penetration distance is shorter, and the signal attenuation is lower. For these reasons, for the same focal distance, the smaller

the subject is, the better the 3D image quality is.

Figure 4-12 shows the long axis views of the Phantom Small (A) and Medium (B). Although some air bubbles appear in the Picture A, the contour of the Phantom Small is very clear and visible. However, some part of the Phantom Medium is completely lost in Picture B. The quality of B is much worse than the quality of A.

The Phantom Large is about 205 ml, which is more than two times larger than the Phantom Medium (86 ml). After the image acquisition following the same protocol, it turned out that most of the Phantom Large was lost, and the image quality of Phantom Large was too bad to apply the semi-automated algorithms.

The iE33 3D echo system is a commercial echocardiography designed for clinical applications. Even for the dilated human LV, it is rare that the LV volume can reach 200 ml. Therefore, it is possible that the 3D echo transducer is not designed for imaging very large chambers.

4.4.2. Reconfirm the conclusions of Chapter 3

After the statistical analysis of the semi-automated volume measurements of the new phantoms, the accuracy and reproducibility of the 21 algorithms on 3D echo images was assessed again. The results reconfirmed the conclusions of Chapter 3:

- After the manual initialization, the semi-automated algorithms were able to trace the tissue-mimicking phantoms inner wall surfaces and measure volumes quantitatively from 3D echo images.
- The 3D edge operator (G_2: 5×5×5) showed better performance than the two 2D operators ($G_{1.1}$: 5×5 and $G_{1.2}$: 11×11) did in both accuracy as well as the reproducibility.
- The interpolation algorithms significantly improved the volume measurements in accuracy.
- The FEM improved further the surface tracing and volume measurements in accuracy.

Different from the balloon phantom measurements, the ROI initialization method 2

was chosen this time. This is because the inner and outer wall surfaces of tissue-mimicking phantoms have similar edges in the 3D echo images (Figure 4-12). Therefore, more precise ROI was required; otherwise the outer wall surface would be detected too. Six points were selected manually instead of two, but it was still very easy to operate and user friendly. Once the ROI was established, all the 21 algorithms started the border tracing and volume measurements automatically.

As mentioned in Chapter 3, extending the 2D edge operator size can reduce the noise in images, but broaden the apparent width of edges. For quantitative volume measurements, bigger operator size leads to certain underestimation of the volumes. However, the 3D edge operator can reduce the noise with small operator size, and thus it can minimize the edge broadening effect.

Thanks to the precise control of the phantom geometry, this phenomenon was observed in a new perspective. Figure 4-11 shows the comparison of the three algorithms plus FEM. The only difference of the three algorithms was the application of different edge operators ($G_{1.1}$, $G_{1.2}$ and G_2). Both plots of the Phantom Small and Medium indicate that, the green ($G_{1.1}$) and purple curves (G_2) are almost overlapped, but the blue ($G_{1.2}$) curve is under them for most of the slice numbers. It means that $G_{1.2}$ tended to give a smaller volume for measuring the same slice.

For reproducibility, since $G_{1.2}$ and G_2 can reduce the noise better, it is easier for the border detection and interpolation algorithms to give more stable results. Table 4-8 and the ANOVA showed that the reproducibility of $G_{1.2}$ and G_2 was significantly better than the reproducibility of $G_{1.1}$.

Figure 4-13 and Figure 4-14 demonstrate the border tracing (white dots) of the Phantom Medium and Phantom Small inner wall surfaces on 3D echo images by G_2B_2, $G_2B_2I_2$ and $G_2B_2I_2$+FEM.

In column 1 of both figures, it shows that the border detection B_2 failed either if the phantom wall was missed, or the contrast of the phantom wall was too low. But in column 2, with the help of I_2, the failed detections were interpolated. This explained why the interpolation algorithms improved the volume measurements.

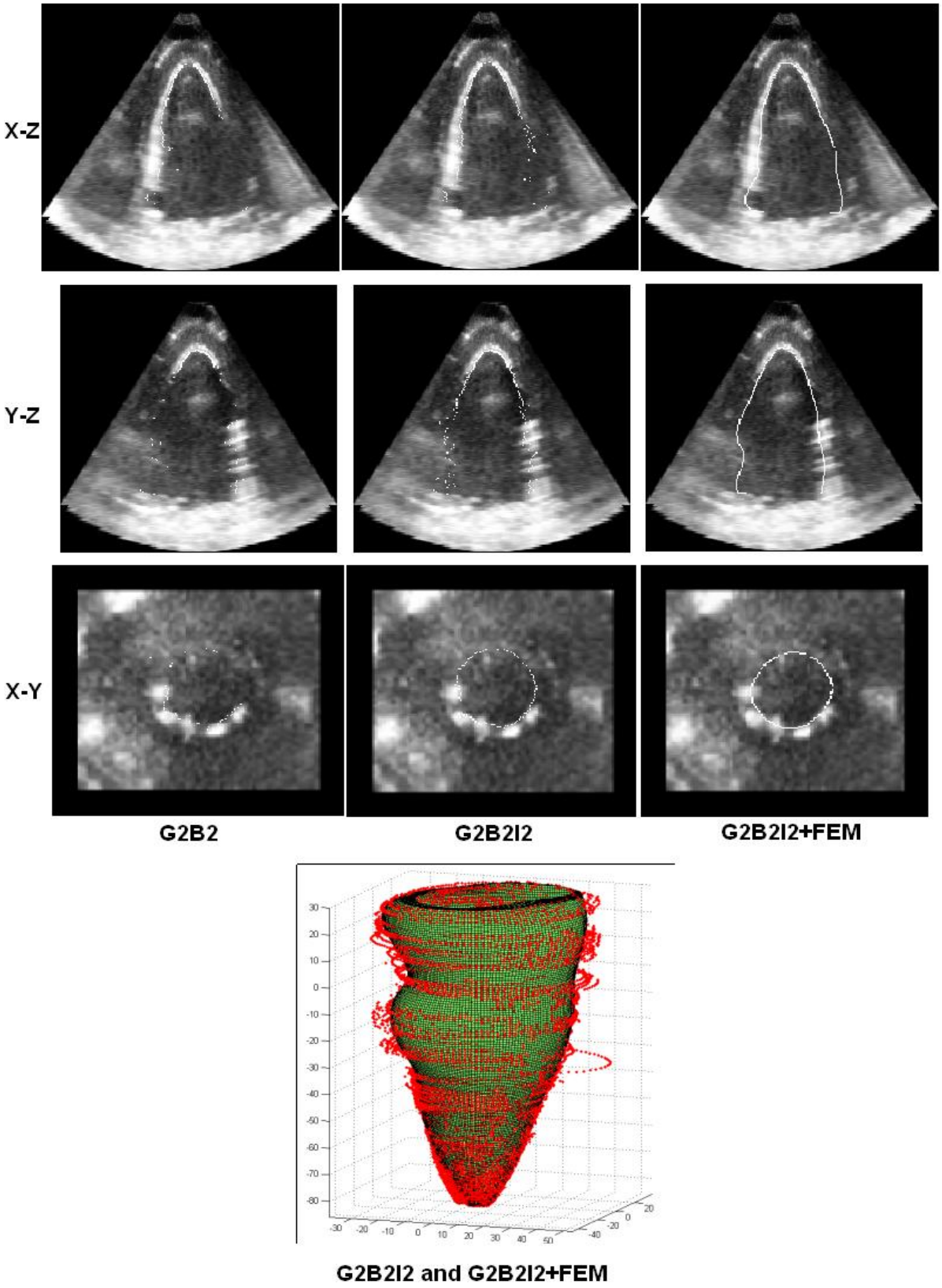


Figure 4-13 Comparison of different algorithms for border tracing (white dots) of the Phantom Medium inner wall surface on 3D echo image: column 1, 2 and 3: G_2B_2, $G_2B_2I_2$ and $G_2B_2I_2$+FEM; row 1, 2 and 3: extracted x-z, y-z and x-y plane. Last picture: overlap of 3D reconstruction: $G_2B_2I_2$ (red dots) and FEM (green meshes)

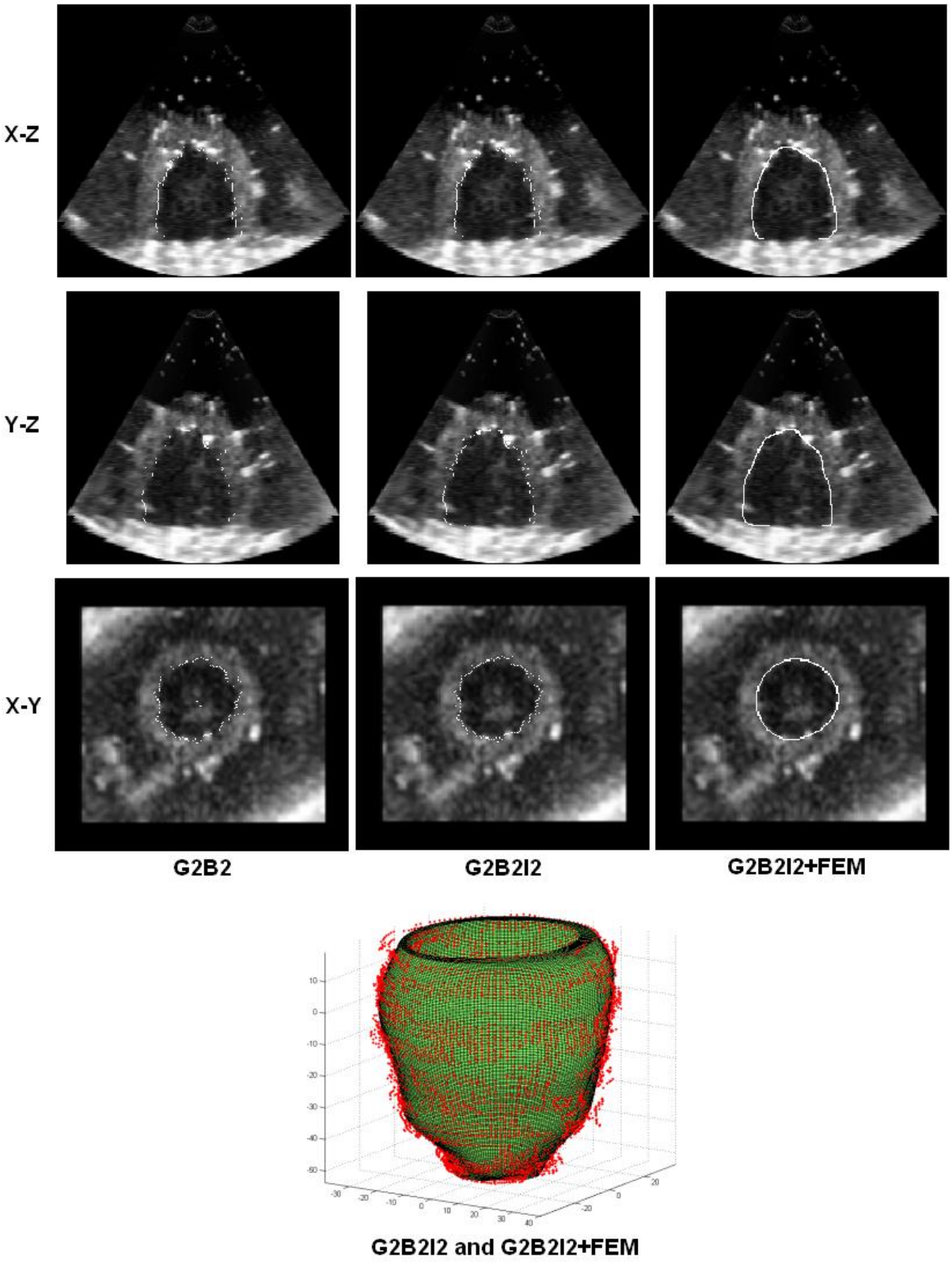


Figure 4-14 Comparison of different algorithms for border tracing (white dots) of the Phantom Small inner wall surface on 3D echo image: column 1, 2 and 3: G_2B_2, $G_2B_2I_2$ and $G_2B_2I_2$+FEM; row 1, 2 and 3: extracted x-z, y-z and x-y plane. Last picture: overlap of 3D reconstruction: $G_2B_2I_2$ (red dots) and FEM (green meshes)

After applying the FEM, in column 3 of both figures, the border tracing was even better. The last pictures show the 3D reconstruction of the inner wall surfaces of the Phantom Medium and Small, given by $G_2B_2I_2$ (red dots) and $G_2B_2I_2$+FEM (green meshes). Obviously, for some slices, the I_2 provided wrong border delineations, because it only considered the information from one 2D slice. But FEM integrated the information from all the slices, thus it improved the surface tracing further. However, because the 3D echo image quality was influenced too much by the noise and artefact, the geometry of the phantoms was distorted and deformed in the images. Therefore, none of the semi-automated algorithms can give a perfect 3D ellipsoidal surface reconstruction.

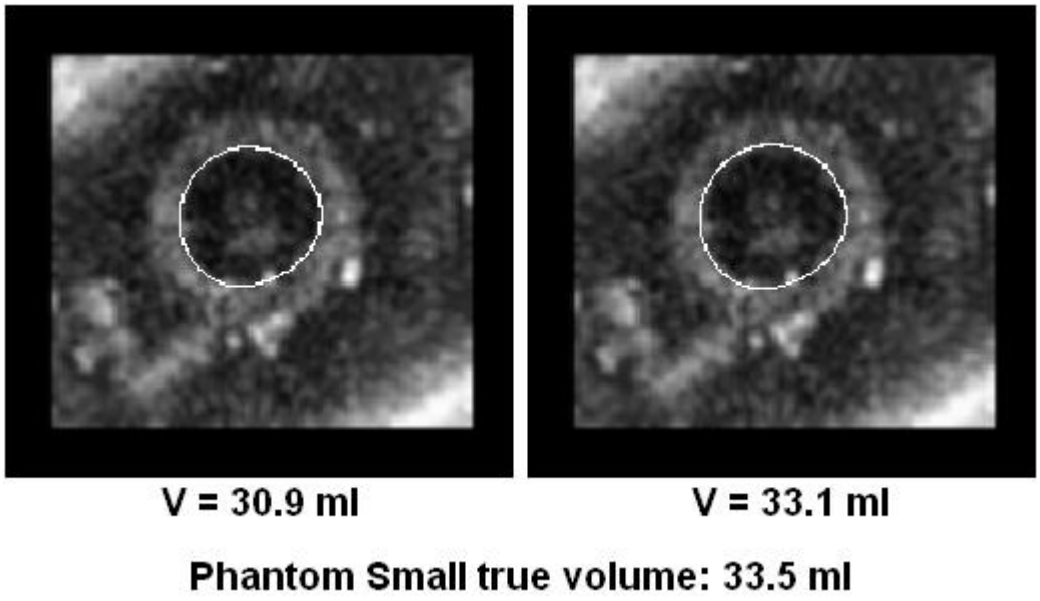


Figure 4-15 Short axis plane extracted from a 3D echo image of Phantom Small: the left picture shows the delineation of the phantom inner wall edge given by $G_2B_2I_2$+FEM. Based on this delineation, the measured volume V=30,9ml; the right picture shows the delineation that extends the original one outwards for one voxel. Based on this delineation, the measured volume V=33.1ml. The actual phantom volume is 33.5ml.

Another thing should be pointed out here is: in Chapter 3, because of the low spatial resolution of the 3D echo images, the true balloon phantom inner wall boundaries were not the balloon inner wall edges in the 3D echo images, but three voxels outward the inner wall edges (Section 3.4.3). In the measurements of the tissue-mimicking phantoms, the same analysis was performed, and it turned out to be that the true phantom inner

wall boundaries were one voxel, instead of three voxels, outwards the inner wall edges in the 3D echo images (Figure 4-15).

Figure 4-15 indicates a very interesting phenomenon. It is difficult to notice the difference between the original delineation (left picture) and the one voxel extended delineation (right picture). However, for all the semi-automated algorithms, extending one voxel resulted in 5% to 7% increase of the measured volumes of both Phantom Small and Medium. This was consistent with the balloon phantom volume measurements (Figure 3-10). All the data demonstrated in this chapter used the one voxel extended delineations for measuring the phantom volumes.

In Chapter 3, delineations of the balloon phantoms were suggested to extend three voxels outwards. However, in this analysis, extending one voxel was enough. The possible explanation is that the newer Philips iE33 3D echo system can provide better imaging quality than the old Sonos 7500 echo system. The edge spread effect was reduced to offer a sharper intensity change. Therefore, the tissue-mimicking phantom inner wall boundary definition was more accurate.

4.4.3. The comparison between 2D and 3D echo images

Comparing with the conventional 2D echo transducer, the 3D echo transducer is much more complicated. 3D echo scans require more computational power for signal processing. The spatial and temporal resolution of 3D echo images was worse than 2D echo images, and 3D images suffer more from the noise and artefact, as shown in Figure 4-7 and Figure 4-8.

For acquired 2D short axis slices of tissue-mimicking phantoms, because of the excellent spatial resolution, the low noise and the free of artefact, semi-automated algorithms provided very accurate volume measurements on 2D echo images. The differences between measured and known slice volumes were small, as shown in Figure 4-9. For 3D images, because of the lower spatial resolution, the high noise and artefact, semi-automated algorithms cannot perform the border tracing as well as they did on 2D images. The differences between measured and known slice volumes were relatively bigger (Figure 4-10) for acquired short axis slices. However, Table 4-4 and Table 4-7

shows that the differences between measured and known phantom volumes were another story. The semi-automated algorithms provided much smaller biases on 3D echo images than they did on 2D images. This is because 3D echo images provided information of the areas around the phantom apex and base, which the conventional 2D echocardiography was not capable of.

In the clinical applications, for quantitative evaluation of patient's LV volumes and EF, it is neither possible to use a 2D echo transducer to scan LV short axis slices with equal distance, nor possible to acquire enough number of short axis slices. The most frequently used clinical method is the bi-plane ellipsoid method (section 1.3.1.). But the drawbacks, such as foreshortened views of LV long axis images, are difficult to avoid, and cause significant underestimation of the LV volumes.

The technology of 3D echocardiography had steady developments in the last decade, and it is still developing. The superior advantage of using 3D echo images for quantitative volume measurements was well demonstrated in this analysis. With the suitable imaging analysis algorithms, the delineation of human LV endocardial surfaces (in this case, phantom inner wall surfaces) can be performed in 3D covering the entire cavity. The bias caused by simple geometric assumption and foreshortening can be eliminated. With the continuous improvement of spatial and temporal resolutions of 3D echo images, more promising measurements should be achieved.

4.4.4. Different performances of B_1 and B_2 on 2D and 3D echo images

After applying the six semi-automated algorithms ($G_{1.1}B_1$, $G_{1.1}B_2$, $G_{1.1}B_1I_1$, $G_{1.1}B_2I_1$, $G_{1.1}B_1I_2$ and $G_{1.1}B_2I_2$) on 2D echo images, Table 4-3 shows that the biases of algorithms with B_1 (Table 4-3, column 1, 3 and 5) were smaller than the biases of algorithms with B_2 (Table 4-3, column 2, 4 and 6) respectively. With or without interpolations (I_1 or I_2), B_1 performed better border delineations and volume measurements than B_2 did on 2D echo images.

The same six algorithms were also applied on 3D echo images. However, Table 4-6 shows that the biases of $G_{1.1}B_2I_1$ and $G_{1.1}B_2I_2$ were smaller than the biases of $G_{1.1}B_1I_1$ and $G_{1.1}B_1I_2$ respectively, for measuring both phantoms. This result was not the same as

in Table 4-3. So why did B_1 offer more accurate measurements on 2D echo images, whereas B_2 performed better on 3D echo images?

As introduced in Chapter 2, B_1 always considers the strongest edges as the wall borders, but B_2 estimates the possibility of different edges being the wall borders. Only if the possibility is high enough, the corresponding edge is considered as the wall border by B_2.

The 2D images have low noise and high contrast of the phantom wall boundary. The strongest edges were indeed the phantom wall borders in most cases. Therefore, B_1 can detect most of the wall borders correctly. Different from B_1, B_2 increases the specificity of the border detection. It rejected some of the detected edges because it considered that the possibility of these edges being the true wall borders is not high enough. Because of this reason, B_1 showed generally better accuracy than B_2 on 2D echo measurements.

The 3D images have high noise and low contrast. Many of the detected edges by B_1 were actually the noise rather than the true phantom wall borders. These false-positive border detections confound the performance of the interpolation algorithms. Unlike B_1, B_2 increased its specificity. The detected edges by B_2 were mostly the true phantom wall borders. Thus the interpolation algorithms were not likely to be misled. In other words, comparing with B_1, B_2 was better with interpolation algorithms (I_1 or I_2) for border delineations. The algorithms with B_2 and interpolations showed better accuracy in measuring phantom volumes. After applying ANOVA on the measured phantom slice volumes from all the 18 algorithms (including the same six algorithms applied on 2D echo images), it also confirmed that the border detection algorithm B_2 showed significantly better accuracy than B_1.

Therefore, for tracing subject boundaries on high noise and low contrast 3D echo images, B_2 plus interpolation algorithms can provide better delineation and volume measurements.

4.5. Conclusion

Compared with conventional 2D echo images, 3D echo images showed higher noise and artefact. But 3D images can provide more information for the phantom apical and basal areas, so that more accurate quantitative volume measurements were achieved.

The findings of the balloon phantom study were reconfirmed: in order to achieve a better quantitative measurement of 3D echo images, the imaging analysis should be performed from all three dimensions. As this utilized the information contained in 3D images, so that the negative influence of noise and artefact can be maximally minimized, and the accuracy as well as reproducibility in quantitative measurement can be improved.

Border detection algorithm B_1 and B_2 showed different performances on 2D and 3D echo images for measuring tissue-mimicking phantom volumes. For higher noise 3D echo images, B_2 was better than B_1.

To sum up, for surface tracing and volume measurements of tissue-mimicking phantoms on 3D echo images, 3D edge operator G_2 + border detection B_2 + interpolation algorithm I_2 + Finite element model provided the best accuracy and reproducibility.

In the next chapter, an algorithm validation study was performed. Clinical Cardiac Magnetic Resonance (CMR) images were acquired. The border detection and interpolation algorithms were applied on the images to trace human LV endocardial surfaces and measure LV volumes. The results were compared with the gold standard manual tracing measurements.

Chapter 5. Validation of the semi-automatic algorithms in cardiac magnetic resonance images

In Chapter 3 and 4, the newly developed semi-automatic algorithms were tested and compared by performing quantitative volume measurements on echo images of balloon phantoms and tissue-mimicking phantoms. The benefit of using the laboratory phantoms is that, with precisely controlled geometry as well as acoustic property, the reference objects are reliable and trustworthy. It is not only the performance of the border tracing and volume calculation algorithms, but also the function of the 2D and 3D echocardiography systems that can be accessed objectively via these references. The drawback is that, compared with the laboratory phantoms, the morphology of real human hearts is much more complicated and also dynamic. The clinical environment is very different from the laboratory environment, so the imaging quality varies a lot in the clinic. Therefore, although semi-automatic algorithms were able to handle the laboratory images of phantoms, it is still questionable whether they are able to handle the clinical human images.

Cardiac Magnetic Resonance Imaging (CMRI) has proved to be an accurate and reproducible imaging modality for quantitative analysis of left ventricular function (70, 71). The manual delineation of the endocardial surface of a LV plus the Simpson's Rule on cardiac magnetic resonance images (CMR images) is normally considered as the gold standard for non-invasive quantitative volume measurements of left ventricles (70 to 73). Compared with the echo images, the MR images can provide much better spatial resolution and great contrast between human blood and myocardium, with low level of noise.

In the MRI centre of Newcastle University, a research group is carrying on a long term study about human aging. In their study, investigators need to manually trace the LV endocardial borders slice by slice and calculate the LV volume (Simpson's Rule) of human hearts. Meanwhile, the semi-automatic algorithms were developed to do exactly the same job. Although these algorithms were originally designed to process the echo images with high noise and low spatial resolution, there should be no reason that they

are not able to be used on CMR images.

Despite the aim of this PhD project being mainly about the semi-automatic algorithm assessment on echo images, and comparison between 2D and 3D echocardiography, it was worth performing a validation study in order to compare the semi-automated algorithms, specifically the border detection and interpolation algorithms, with the gold standard (manual tracing on CMRI) on the human subjects. The aim of this particular validation study was to answer the question: with low noise, good contrast and clinical environment human CMR images, whether the border detection and interpolation algorithms can show a comparable performance to the manual technique on quantitative LV volume measurement.

5.1. Imaging acquisition

18 subjects (6 males and 12 females; age range: 21 to 71 years; mean age: 56 years) without any cardiovascular disease participated in this study. Data were collected using a 3T Intera Achieva scanner (Philips, Best, NL). A dedicated 6-channel cardiac coil (Philips, Best, NL) was used with the subjects in a supine position and ECG gating (Philips vectorcardiogram, VCG system). Ethical approval was obtained and signed consent obtained for these human studies.

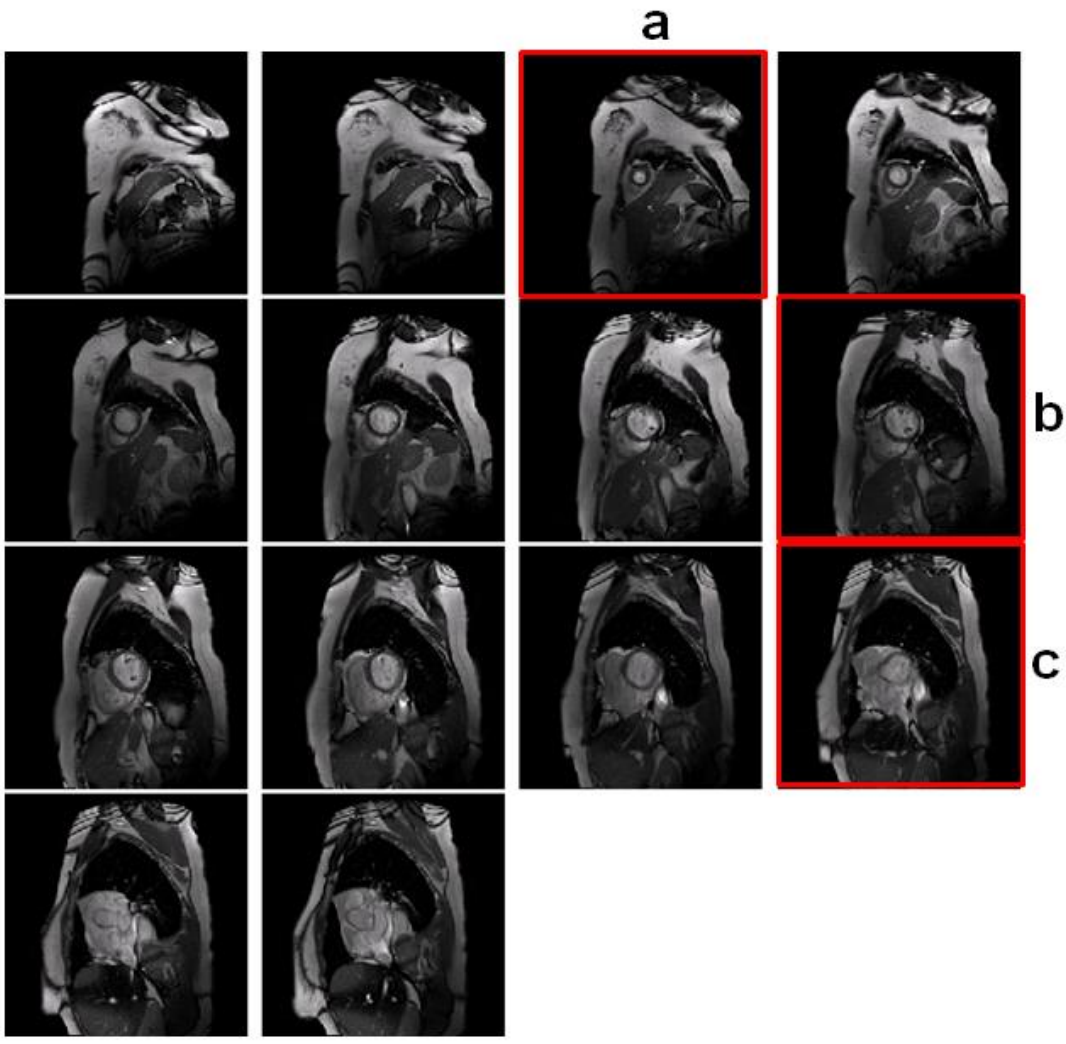


Figure 5-1 Demonstration of a stack of 2D short axis slices (observation sequence: from left to right, top to bottom); Slice a, b and c were selected to establish the ROI

A stack of balanced steady-state free precession images was obtained in the short axis view covering the entire left ventricle (field of view = 350mm, repetition time/echo time

= 3.7/1.9ms, turbo factor 17, flip angle 40^{o}, slice thickness 8mm, 0mm gap, 14 slices, 25 phases, resolution 1.37mm, temporal duration approximately 40ms per phase, dependent on heart rate). An example of the short axis MR images from one subject at LVED is shown in Figure 5-1. Compared with the echo images, those images have low noise and very good contrast between blood, myocardium and other surrounding tissues and organs.

5.2. Imaging analysis

Manual and new semi-automatic imaging analysis techniques were applied for measuring the left ventricular volume parameters of each subject. The manual tracing of the LV endocardial surface was performed by an experienced cardiologist.

5.2.1. Manual measurement

Manual analysis was performed using the cardiac analysis package of the ViewForum workstation (Philips, Best, NL). Manual tracing of the endocardial borders was performed on the short axis slices at end-systole and end-diastole. The contours were reviewed by viewing the cine data with the contours attached. The apical slice was defined as the last slice showing intra-cavity blood pool, as shown in Figure 5-1 a. The basal slice selected for analysis for end-diastole and for end-systole occurred when at least fifty percent of the blood volume was surrounded by the myocardium (74), as shown in Figure 5-1 c. Papillary muscles were included from calculations of LV volumes. Left ventricular end-systolic volumes (LVESV), end-diastolic volumes (LVEDV), stroke volumes (LVSV=LVEDV-LVESV) and ejection fraction (LVEF=LVSV/LVEDV) were calculated by Simpson's rule. The LVEDV was measured after mitral valve closure, and the LV cavity reached its maximum volume. The LVESV was measured on the images with the smallest LV cavity

5.2.2. Semi-automated measurement

Seven semi-automated analyses ($G_{1.1}B_1$, $G_{1.1}B_2$, $G_{1.1}B_1I_1$, $G_{1.1}B_2I_1$, $G_{1.1}B_1I_2$, $G_{1.1}B_2I_2$ and $G_{1.1}B_2I_2$+FEM) were performed on the short axis slices selected using the same

protocol as the manual analysis. To select the region of interest (ROI), the ROI initialization method 2 was applied (Chapter2: Section 2.1): Three slices (left ventricle apical, middle and basal slice, Figure 5-1 a, b and c) were selected for ROI initialization. One point in the left ventricular blood pool and one point in the left ventricular myocardial were manually selected for each of the three slices, so that the ROI was established automatically in all slices (Figure 2-2). Once the ROI was established, the seven semi-automated algorithms would start endocardial surface tracing and volume calculation by Simpson's rule.

Despite the excellent imaging resolution and contrast, the MRI can not provide a 3D imaging dataset as the 3D echocardiography can. The 2D short axis slices acquired by MRI have significant slice thickness (8mm). Therefore, the 3D gradient operator G_2 is inapplicable on these images. Fortunately, since the noise in MRI is very low, it is not necessary to apply the 3D edge operator for noise reduction. A simple 5×5 2D gradient operator ($G_{1.1}$) is sufficient in this experiment. Because all the semi-automated algorithms use the same edge operator, it is actually comparing the border detection and interpolation algorithms with the manual technique. For the rest of this chapter, each algorithm is coded without the edge operator $G_{1.1}$ for simplification, as it was used for all images. Similar to the previous laboratory experiments, the traced endocardial surface dataset from B_2I_2 was used to apply the finite element model (FEM).

5.2.3. Statistical analysis

After measuring the left ventricular EDV, ESV, SV and EF, the relationship between each semi-automated algorithm and the manual technique was evaluated by linear regression analysis with correlation coefficient r^2. The agreement was expressed according to the Bland Altman method (67) by calculating the bias (mean difference), ±standard error of the mean difference and ±2×standard deviation of the difference (upper and lower limits of agreement). The significance of the biases was tested by paired t tests with a two-tailed distribution. $P < 0.05$ was considered significant.

5.3. Results

The seven semi-automatic algorithms (B_1, B_2, B_1I_1, B_2I_1, B_1I_2, B_2I_2 and B_2I_2+FEM) measured the volumes for all 18 subjects at both LVED and LVES successfully. It took about 10 seconds to measure one volume, including the manual initialization of the ROI, semi-automatically, and about 10 minutes to measure manually.

5.3.1. Results of linear regression analysis

The values of correlation coefficient r^2 between each semi-automated and manual measurement of EDV, ESV, SV and EF are listed in Table 5-1, and their linear regression analyses are plotted in Figure 5-2, Figure 5-3, Figure 5-4 and Figure 5-5, respectively. Generally speaking, all semi-automated measurements correlated highly with the manual measurements of EDV, ESV and SV, but relatively lowly with EF. This was because there was a progression of performance with EDV being the best and ESV being slightly worse (Table 5-1 first and second row), which lead to the rather poor correlation in calculated parameter EF.

Table 5-1 Correlation coefficient r^2 between each semi-automated and the manual measurement of EDV, ESV, SV and EF

	Correlation Coefficient r^2						
	B_1	**B_2**	**B_1I_1**	**B_2I_1**	**B_1I_2**	**B_2I_2**	**FEM + B_2I_2**
EDV	0.97	0.91	0.98	0.99	0.98	0.99	0.99
ESV	0.90	0.83	0.94	0.94	0.96	0.96	0.95
SV	0.85	0.77	0.90	0.94	0.91	0.95	0.95
EF	0.21	0.17	0.57	0.71	0.58	0.75	0.71

The correlation of measuring EDV was the best for all seven semi-automated algorithms (Table 5-1 the first row), and algorithm B_2I_2 showed the best correlation for all measurements of EDV, ESV, SV and EF (Table 5-1 the sixth column, B_2I_2). Without interpolation algorithm, the border detection algorithm B_1 showed better correlation than B_2 did in all the measurements (Table 5-1, the first and second column). However, after applying border interpolation algorithms (I_1 or I_2), there were obvious improvements of all the correlations, and B_2+I_1 or +I_2 showed the same or even better results than B_1+I_1 or +I_2 did, especially in the result of EF (Table 5-1 the last row).

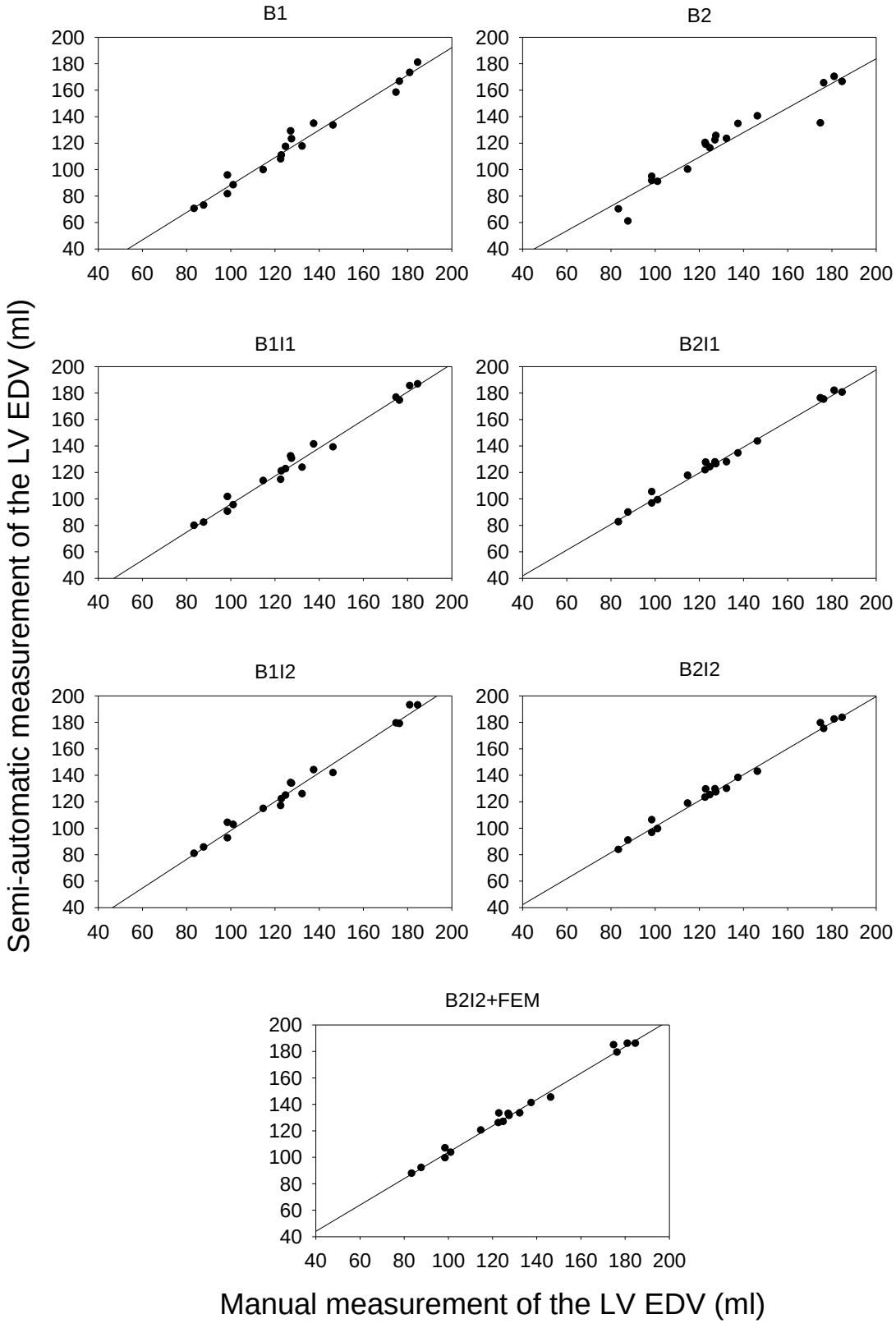


Figure 5-2 Linear regression analysis of each semi-automated measurement against gold standard manual measurement of LV EDV

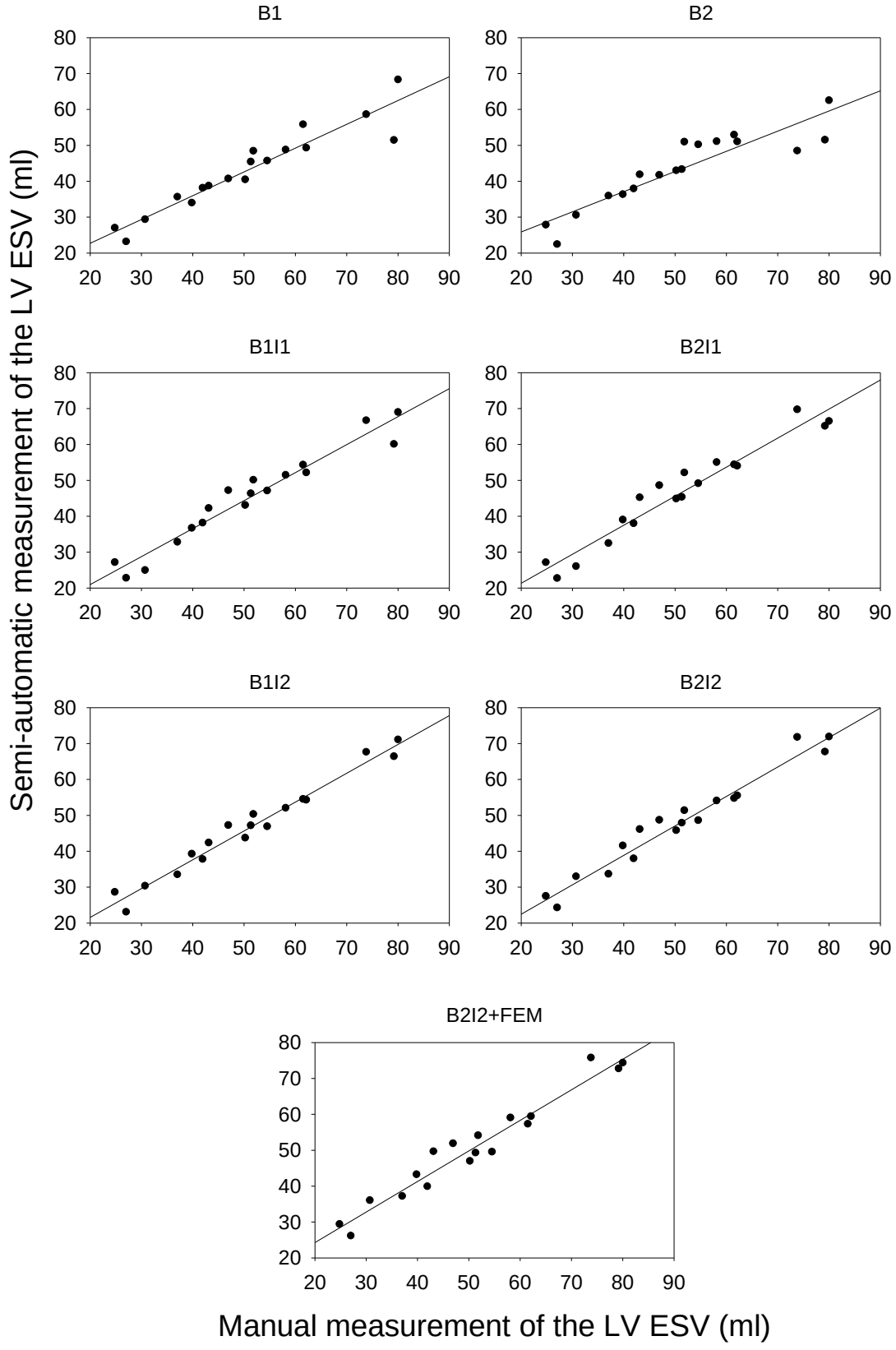


Figure 5-3 Linear regression analysis of each semi-automated measurement against gold standard manual measurement of LV ESV

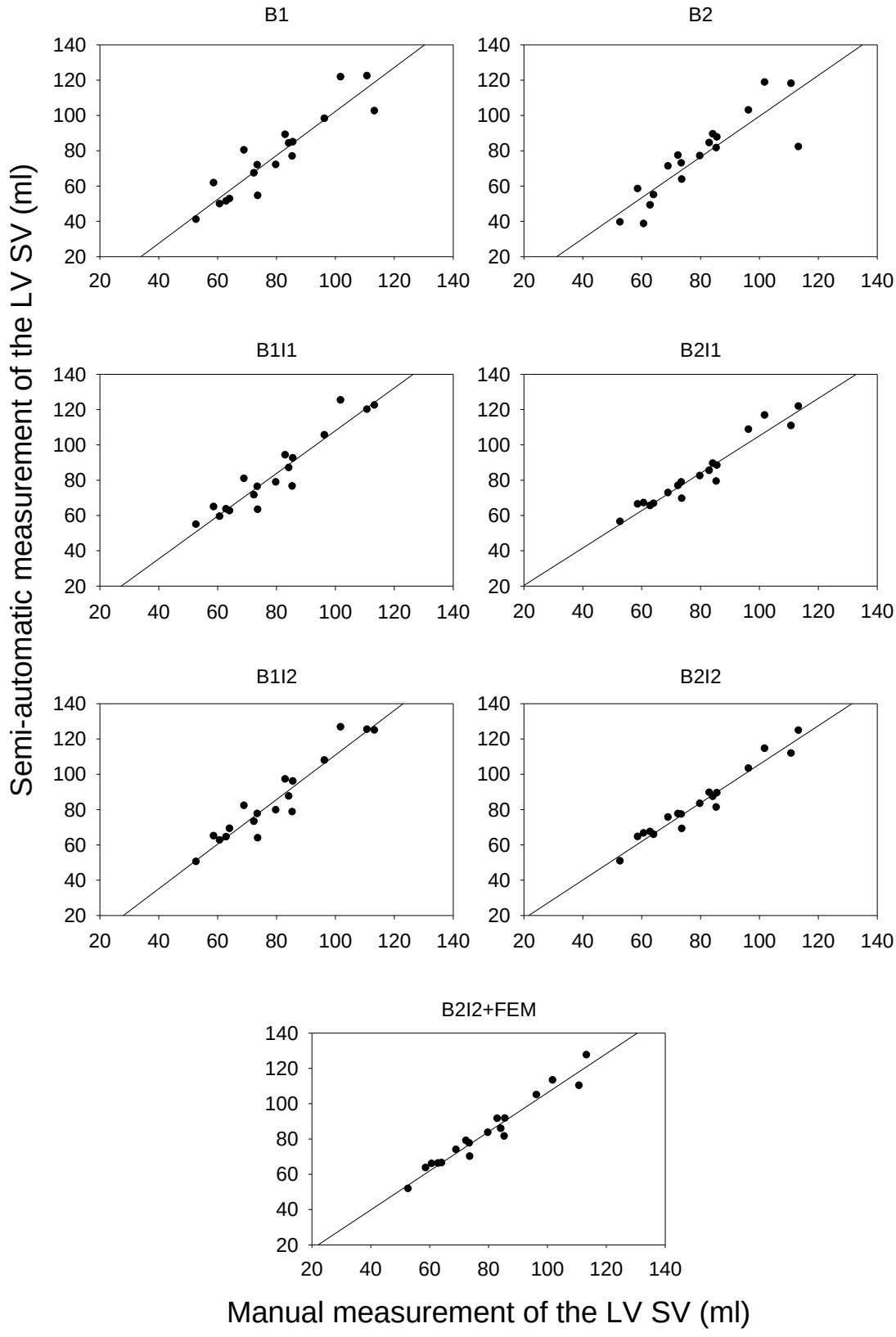


Figure 5-4 Linear regression analysis of each semi-automated measurement against gold standard manual measurement of LV SV

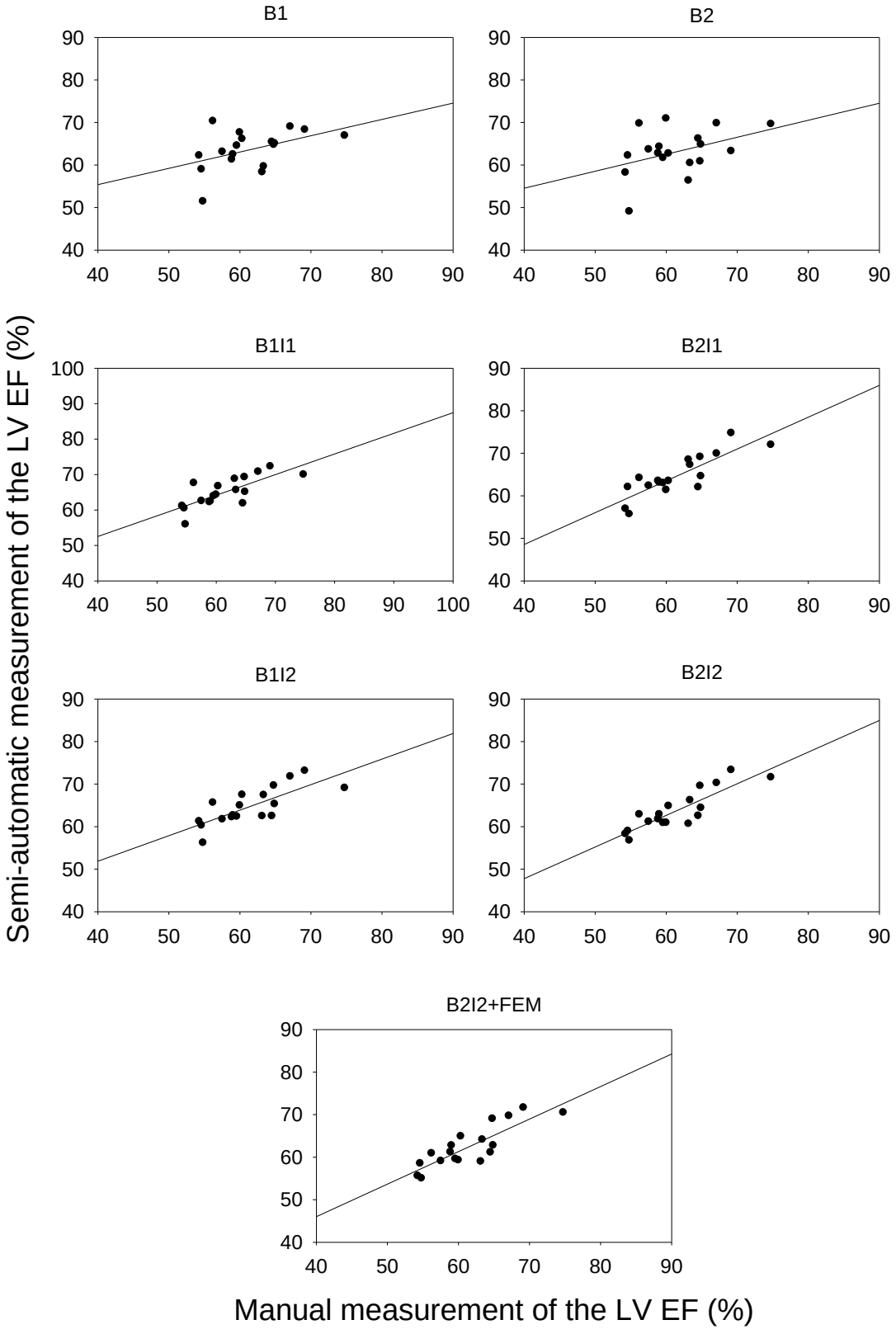


Figure 5-5 Linear regression analysis of each semi-automated measurement against gold standard manual measurement of LV EF

Compared with B_2I_2, the results of correlation did not appear to show an improvement after using the FEM (Table 5-1 the sixth and seventh column). For measuring ESV and EF, the correlation of FEM was even poorer than the correlation of B_2I_2.

5.3.2. Results of Bland Altman analysis

The Bland Altman comparisons of measuring EDV, ESV, SV and EF are plotted in Figure 5-6, Figure 5-7, Figure 5-8 and Figure 5-9. The bias (mean of the difference) ±SD of semi-automated and manual measurements are listed in Table 5-2.

Table 5-2 Bland Altman comparisons (bias±SD) between semi-automated and manual measurement of EDV, ESV, SV and EF

	Semi-automated - manual measurement						
	B_1	**B_2**	**B_1I_1**	**B_2I_1**	**B_1I_2**	**B_2I_2**	**B_2I_2 FEM**
EDV (ml)	-10.4±5.5*	-11.2±9.3*	-2.1±4.6	-0.7±2.9	1.1±5.4	0.7±3.0	3.7±3.0*
ESV (ml)	-7.7±6.5*	-7.7±8.1*	-5.8±4.7*	-4.5±4.5*	-4.5±3.9*	-3.1±4.0*	-0.3±4.0
SV (ml)	-2.8±9.9	-3.6±11.2	3.7±7.8	3.9±4.9*	5.6±8.3*	3.8±4.4*	4.0±4.7*
EF (%)	2.1±5.2	1.6±5.7	3.5±3.5*	3.2±2.9*	3.3±3.5*	2.3±2.7*	1.0±2.9

Note: * significant difference.

Compared with the manual measurements, border detection algorithm B_1 and B_2 showed similar biases, but the standard deviation of B_2 was bigger (Table 5-2 the first and second column). Both B_1 and B_2 significantly underestimated the EDV and ESV.

After using interpolation algorithms (I_1 or I_2), the biases of measuring EDV became much smaller, and there were no significant differences with the manual measurements anymore. Similarly, after using I_1 or I_2, the biases of measuring ESV also became smaller, but the improvement was not as good as for measuring EDV. There were still significant underestimations of the ESV.

Since both B_1 and B_2 showed significant negative biases in measuring EDV and ESV, and SV=EDV-ESV, EF=SV/EDV, the biases of B_1 and B_2 in measuring SV and EF were not significant. After applying I_1 or I_2, however, the biases in measuring SV and EF became significant, because the improvement of measuring ESV was not equally good as measuring EDV.

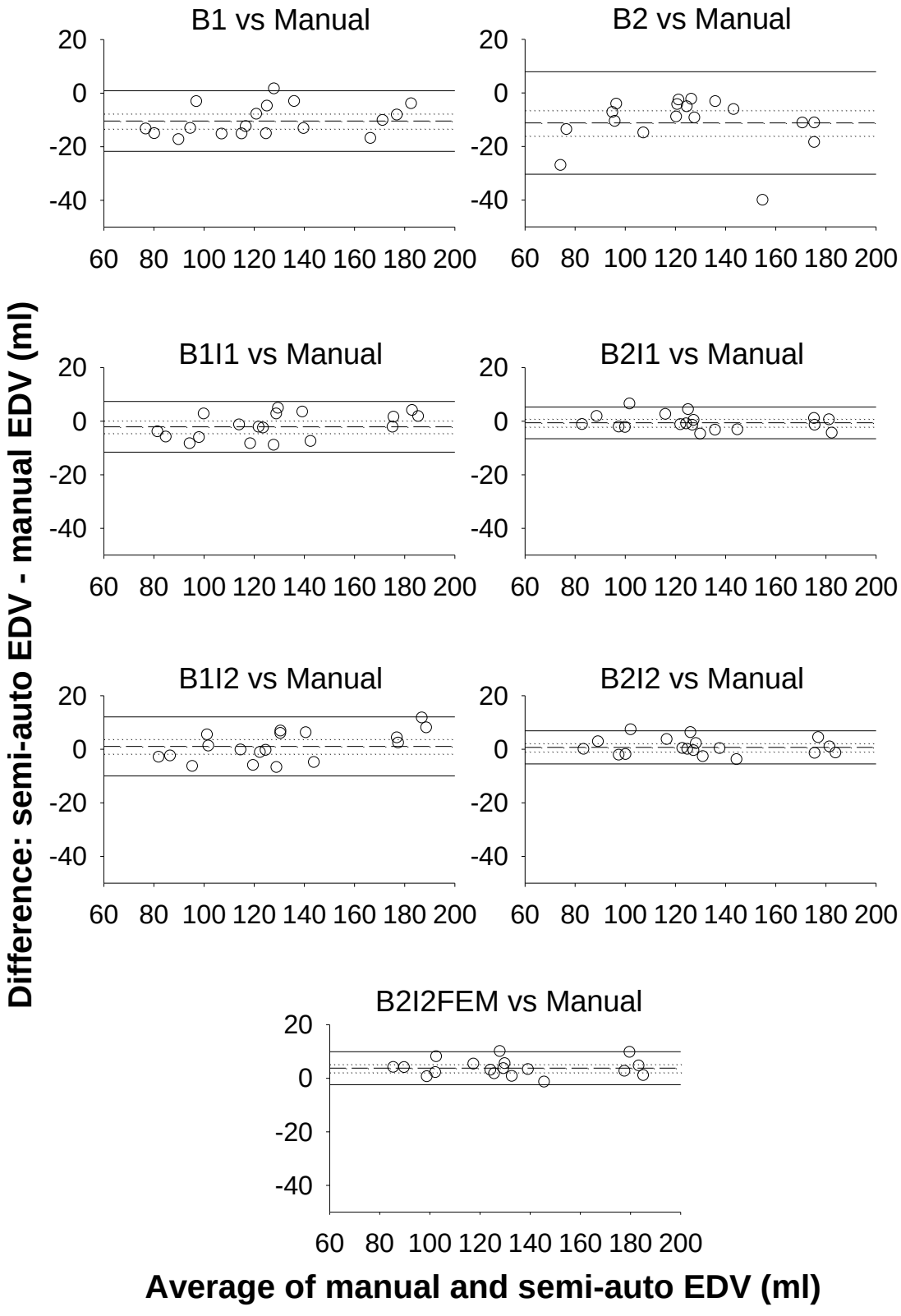


Figure 5-6 Bland Altman plots: Comparison between semi-automated and manual measurement (gold standard) of measuring LVEDV (ml). Dash lines: bias (mean difference); dotted lines: 95% confidence interval (CI) of bias; dash lines: upper and lower limits of agreement

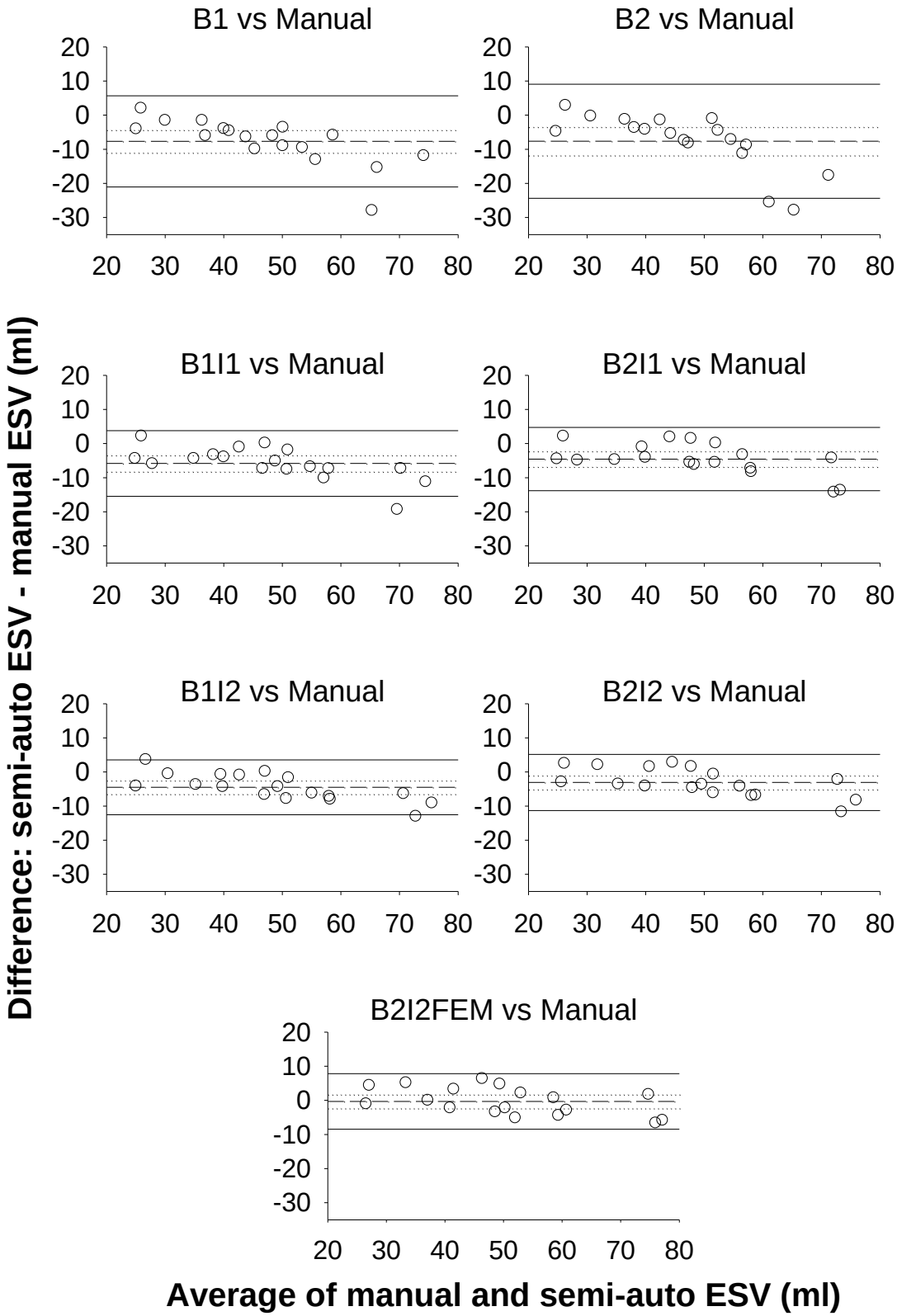


Figure 5-7 Bland Altman plots: Comparison between semi-automated and manual measurement (gold standard) of measuring LVESV (ml). Dash lines: bias (mean difference); dotted lines: 95% confidence interval (CI) of bias; dash lines: upper and lower limits of agreement

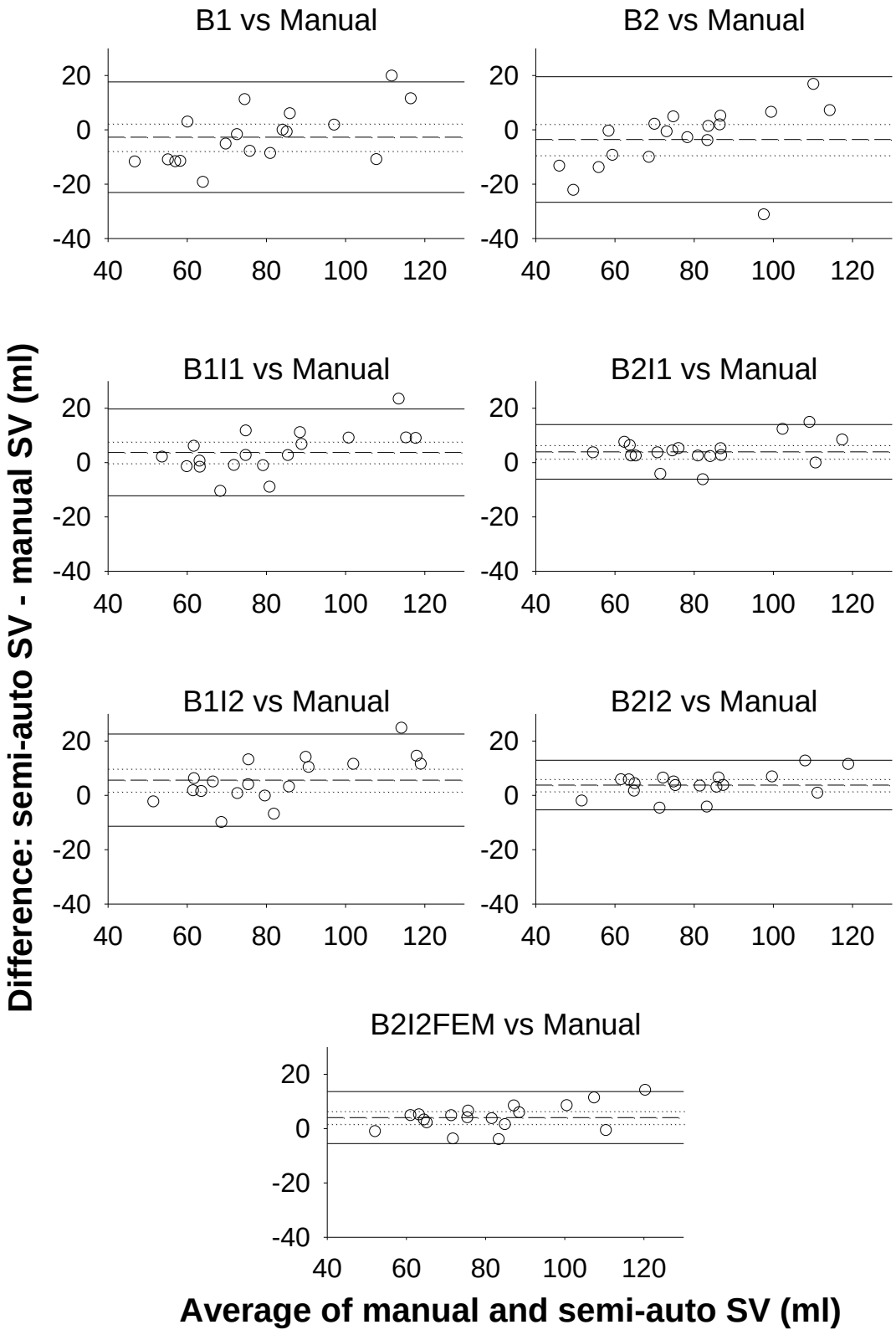


Figure 5-8 Bland Altman plots: Comparison between semi-automated and manual measurement (gold standard) of measuring LVSV (ml). SV = EDV-ESV. Dash lines: bias (mean difference); dotted lines: 95% confidence interval (CI) of bias; dash lines: upper and lower limits of agreement

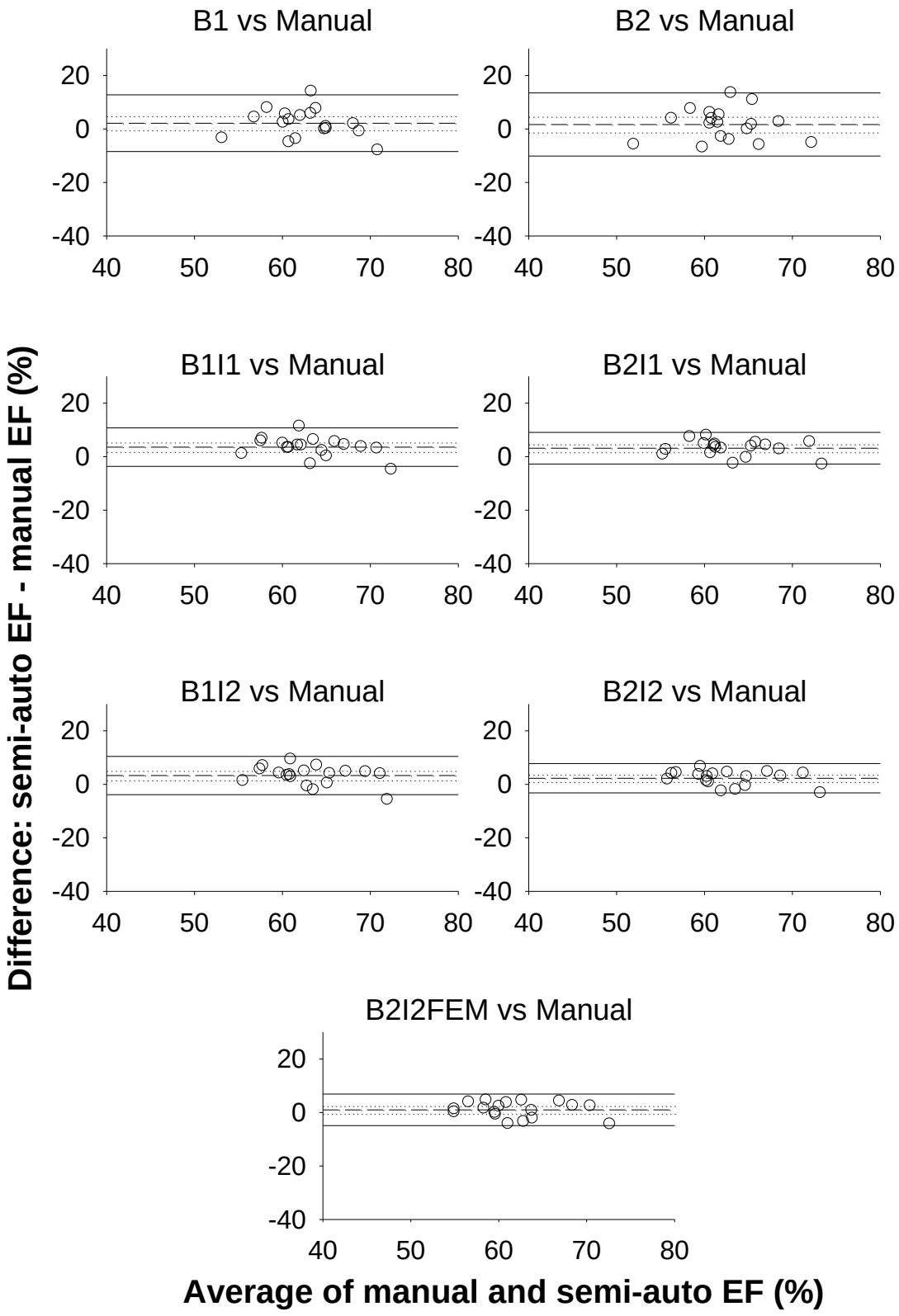


Figure 5-9 Bland Altman plots: Comparison between semi-automated and manual measurement (gold standard) of measuring LVEF (%). EF = SV/EDV. Dash lines: bias (mean difference); dotted lines: 95% confidence interval (CI) of bias; dash lines: upper and lower limits of agreement

After applying the FEM on B_2I_2, the results of measuring EDV and ESV increased about 3ml for both. This resulted in a significant difference in measuring EDV and SV, but a non-significant difference in measuring ESV and EF (Table 5-2 last column).

5.4. Discussion

As the gold standard, the CMRI can provide excellent imaging contrast, superior ventricular endocardial surface, papillary muscle and chordae tendinae definition, which are impossible to achieve by the RT3DE. Thus, these advantages sometimes cause extra problems for the semi-automated algorithms that were originally designed for RT3DE images.

5.4.1. What is the problem of B_1 and B_2?

Without interpolation algorithms, border detection algorithm B_1 provided better correlation with the standard, smaller bias in measuring EDV, as well as smaller standard deviation in measuring EDV, ESV, SV and EF. However, did B_1 provide better border tracing than B_2 did?

As it has been explained in Chapter 2, B_1 and B_2 follow different principles for edge detection. B_1 considers the strongest edge as the endocardial border. If B_1 finds two or more edges equally sharp, the detection along that radial line would be considered as failed. B_2 tries to evaluate all the detected edges, and considers the one with the highest probability as the endocardial border. If B_2 cannot decide which edge is the endocardial border, the detection along that radial line would be considered as failed. Because of this difference, the tracing of one same CMR image sometimes showed significantly different delineations.

Figure 5-10 demonstrates an example of the border tracing by seven semi-automated algorithms on the same 2D short axis CMR image at LVED (basal ventricular slice). Because of the superior imaging contrast and resolution of CMR, in the first picture, the original image, the papillary muscles were captured inside the LV cavity. This is normally difficult to see in RT3DE images. Outside the heart, the lung, fat tissue and

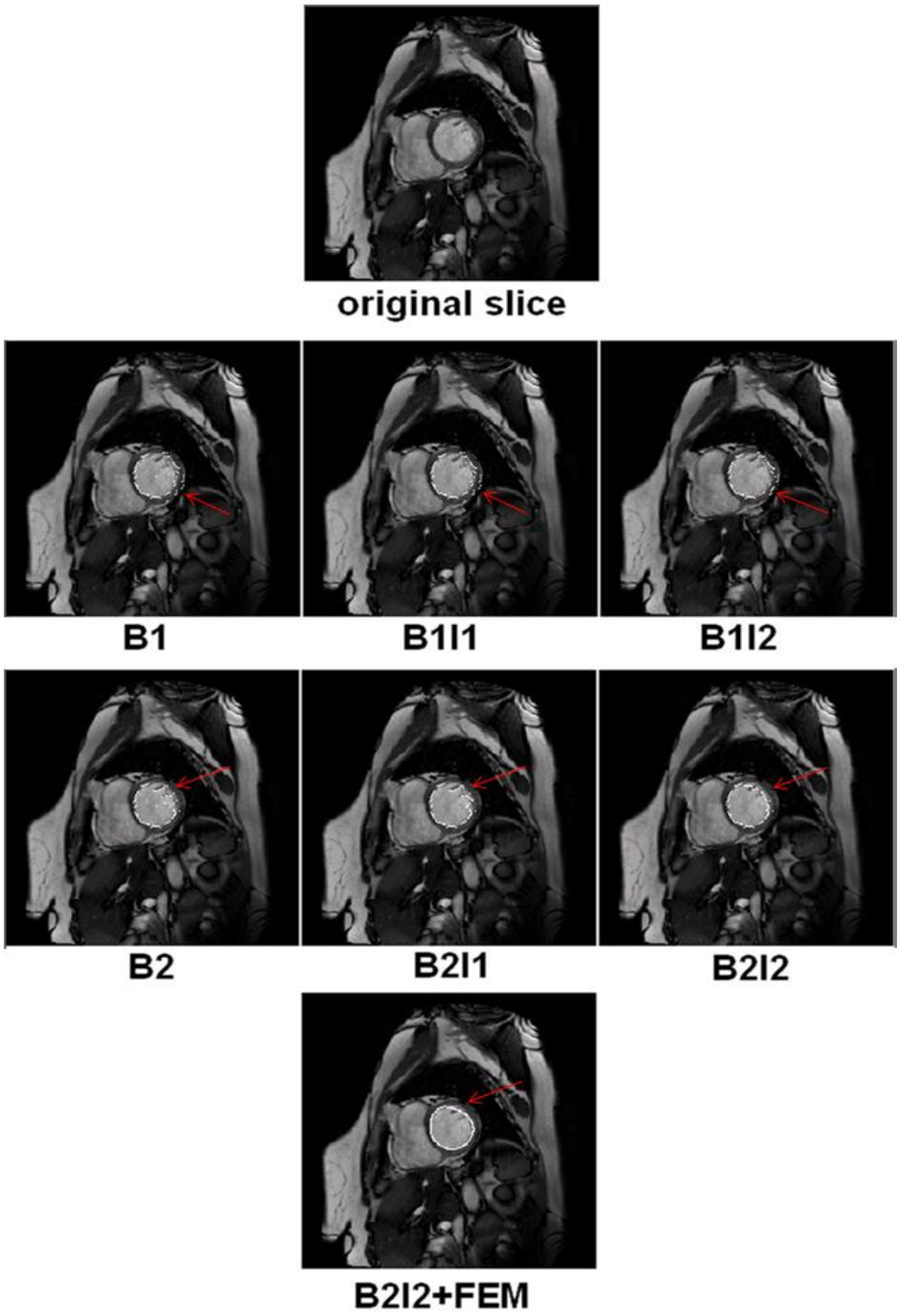


Figure 5-10 Seven semi-automated delineations (white dots) of the endocardial border on a basal-ventricular short axis image at ED

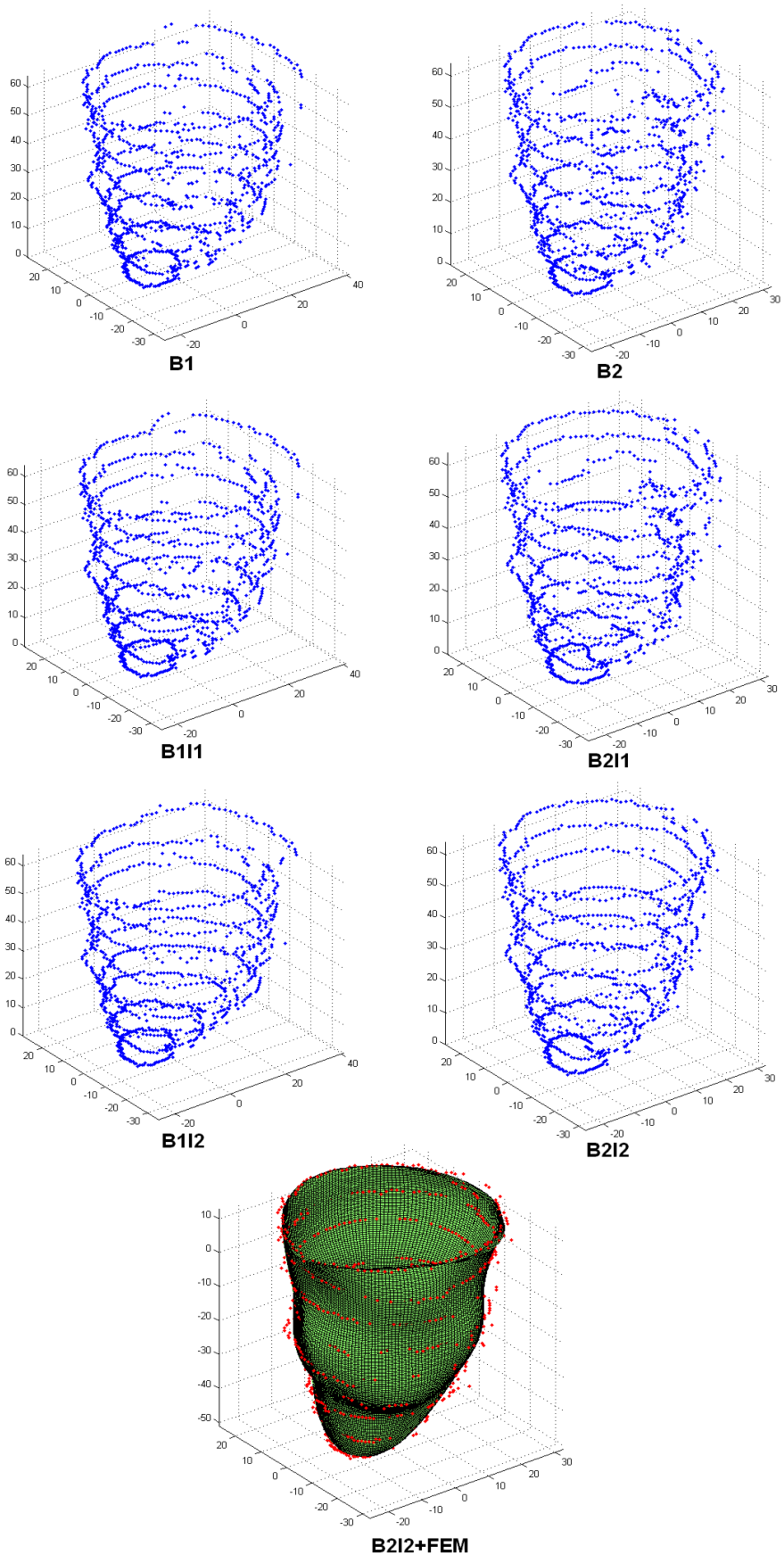


Figure 5-11 3D reconstruction of one LV endocardial surface at ED by seven semi-automated algorithms

other organs surround the heart were also imaged in black or middle grey. This is also impossible to see in RT3DE images. For B_1 and B_2, the new features caused some difficulties in border detection.

The delineation (white dots) in Figure 5-10 (B_1) showed that, in the inferolateral (red arrow) part, B_1 treated the edges from the fat tissue (light gray) to the lung (black) as the LV endocardial borders. This is because these edges are much stronger than the edges from blood pool (light gray) to myocardium (middle gray). In Figure 5-11 (B_1), it can be seen that in the delineation of the basal slice, there was obvious discontinuous tracing of the endocardial wall border. However, B_2 can overcome this problem, as shown in Figure 5-10 (B_2) and Figure 5-11 (B_2). In the same inferolateral part, B_2 succeeded in tracing the endocardial borders, because B_2 is able to distinguish the edges caused by the myocardium or by the lung.

Around the anterior and anterolateral part of the LV in Figure 5-10 (B_1) and (B_2), where the papillary muscle appeared, it can be found that neither B_1 nor B_2 were able to distinguish the papillary muscle and the myocardium perfectly. Figure 5-10 (B_2) showed that, in the area with papillary muscles (red arrow), B_2 treated the tip of the papillary muscle as the endocardial borders in most of the detections. This is because the imaging intensity of the papillary muscle is actually the same as the myocardium (middle grey) in CMR images.

Although both B_1 and B_2 cannot trace the endocardial borders perfectly, the delineation of the lung and papillary muscle can, sometimes, compensate each other in volume calculation for B_1, whereas for B_2, there was no such compensation. This explained why the B_1 correlated better with the standard, and offered smaller bias and standard deviation.

5.4.2. Why did B_2 work better with I_1 or I_2?

Compared with B_1, B_2 increases its specificity but decreases its sensitivity. For one short axis slice, both algorithms scan 120 radial lines (Chapter 2). B_1 tries to find the endocardial borders of every radial line, whereas B_2 tries to find these radial lines with high probability of being the endocardial borders, and abandons the other radial lines.

Therefore, the number of incorrect (false positive) tracing from B_1 was more than from B_2. This makes a difference in border interpolation algorithms.

After comparing the delineation in picture (B_1I_1) and (B_1I_2) with (B_2I_1) and (B_2I_2) in Figure 5-10 and Figure 5-11, in can be found that both I_1 and I_2 were trying to improve the delineation in all cases. However, since the number of the incorrect detection from B_1 was much more, the interpolation algorithms also suffered more from them. In picture (B_1I_1) and (B_1I_2) of Figure 5-10, even with the help of I_1 and I_2, the border delineation was still located on the lung edges in the LV inferolateral area (red arrows). In Figure 5-11 (B_1I_1) and (B_1I_2), the differences are also illustrated clearly in the 3D reconstruction.

In picture (B_2I_2) of Figure 5-10, with the help of I_2, it shows that the delineation was very close to perfect. Even the influence of the papillary muscles (red arrow area) was alleviated because of the help from all corrected detections. In Figure 5-11 (B_2I_1) and (B_2I_2), the 3D reconstruction of the LV endocardial surface was reasonable and smooth, especially in picture (B_2I_2).

5.4.3. Why did the interpolation algorithms work better in ED than in ES?

Without interpolation algorithms, both B_1 and B_2 showed significant underestimation of the LV volume at ED and ES. However, after using the interpolation algorithms, there was a significant improvement in measuring the LVEDV, but relatively moderate improvement in measuring the LVESV.

Figure 5-12 and Figure 5-13 illustrate the border tracing and 3D reconstruction of the endocardial surface of the same subject in Figure 5-10 and Figure 5-11 at ES. Different from the short axis slice at ED, the slice at ES showed much thicker myocardial wall, and smaller LV cavity. But what is more important is that the papillary muscle touched the myocardium because of the contraction of the heart muscle, whereas at ED, most of the papillary muscle was disconnected from the myocardial wall. This feature makes the detection of the endocardial border into a more difficult level.

The Figure 5-12 (B_1I_1), (B_1I_2), (B_2I_1) and (B_2I_2) showed that all of the semi-automated algorithms traced the tip of the papillary muscle rather than the real

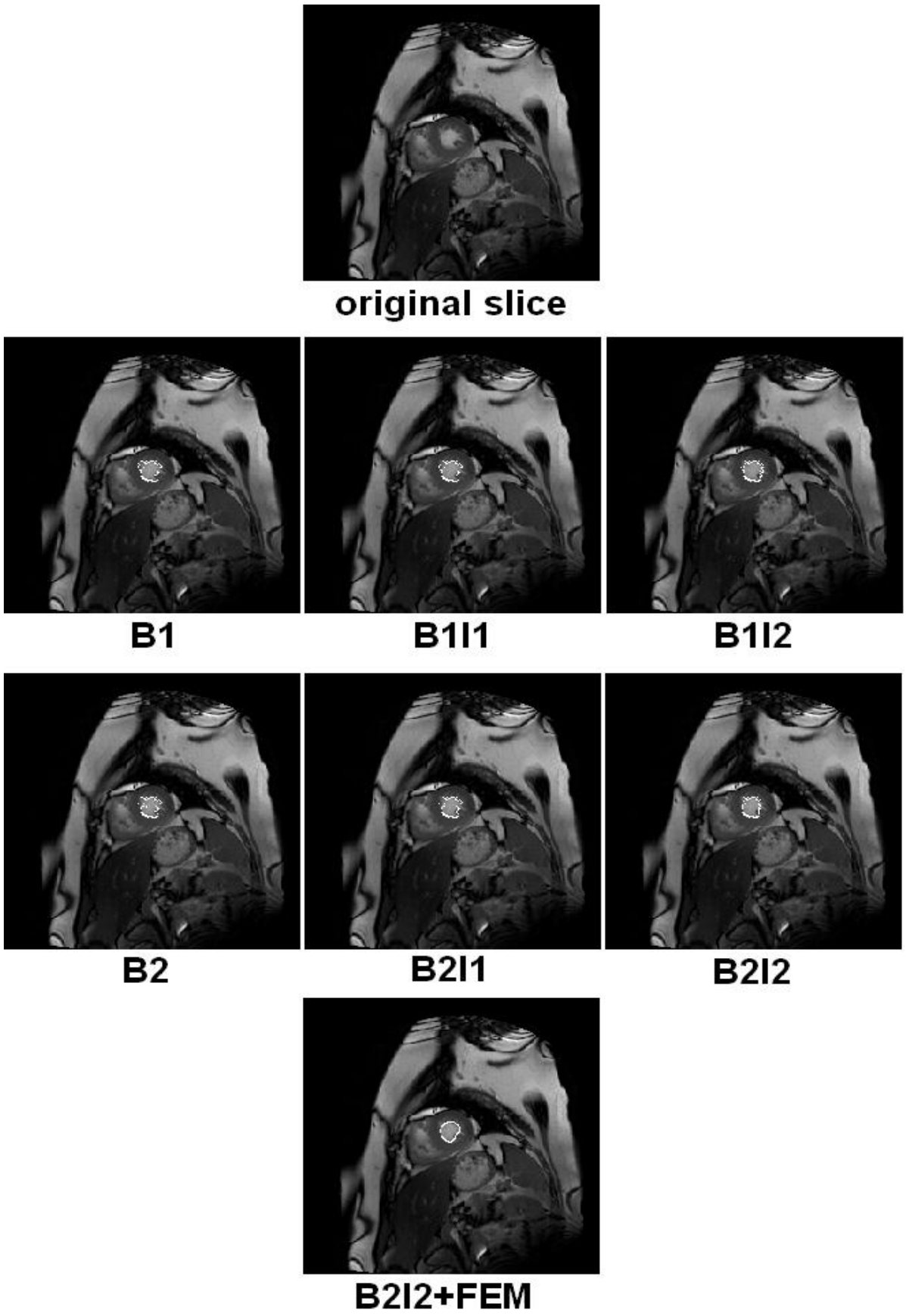


Figure 5-12 Seven semi-automated delineations (white dots) of the endocardial border on a mid-ventricular short axis image at ES

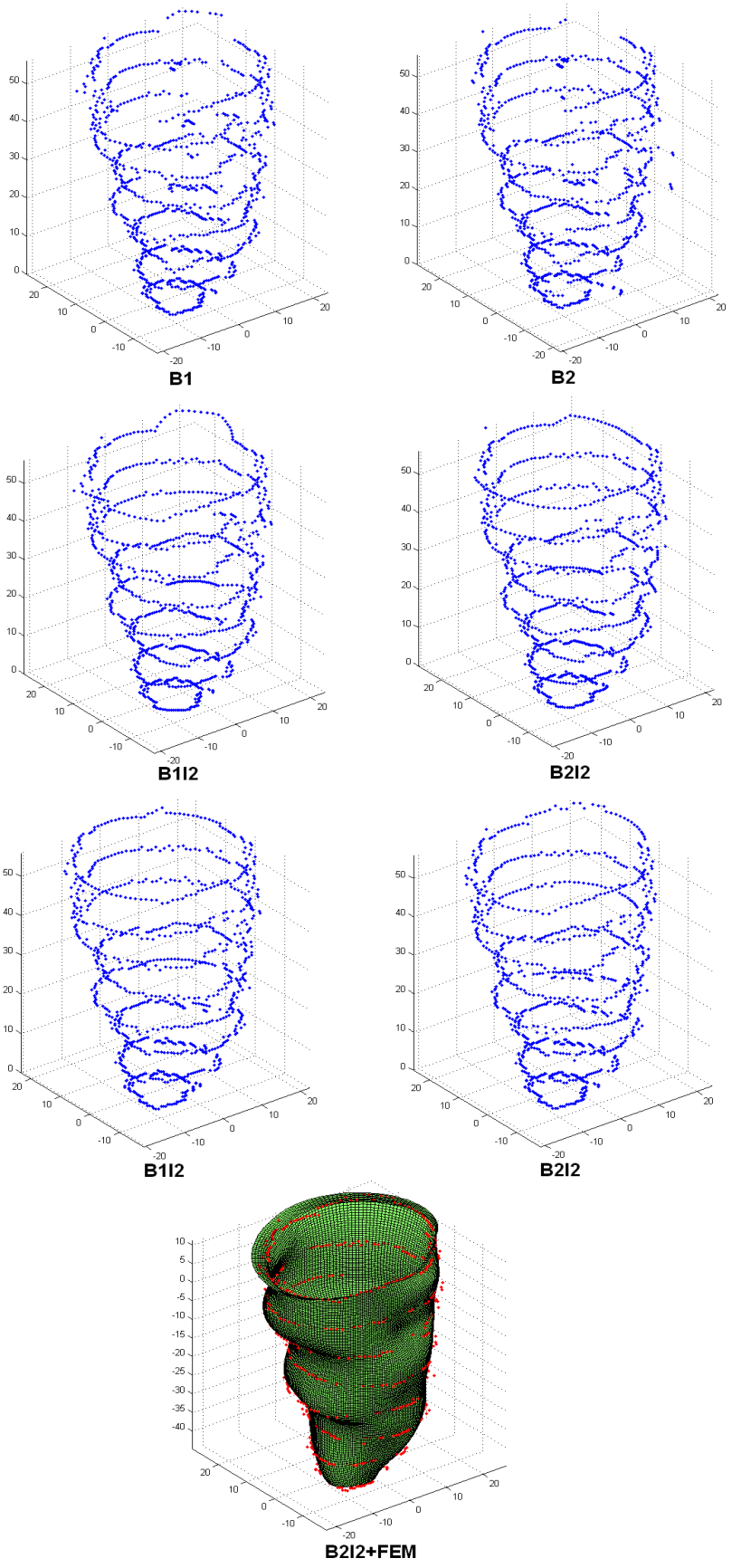


Figure 5-13 3D reconstruction of one LV endocardial surface at ES by seven semi-automated algorithms

endocardial border, because it is impossible to distinguish them only based on the imaging intensity information. However, with the pre-knowledge of the LV morphology and the accumulated experience, the cardiologist is able to delineate the endocardial borders, and separate the papillary muscle at LVES, although with certain level of the inter- and intra-observer variability (75, 76).

Because of this reason, the interpolation algorithms only provided limited improvement of the border tracing at ES, and the bias with the manual method in measuring ESV was still significant for all the semi-automated algorithms.

5.4.4. What is the problem in using FEM?

After using the finite element model on B_2I_2, the correlation and the bias for measuring EDV, ESV, SV and EF did not show significant improvement, and on the contrary, it made the results worse sometimes, compared with B_2I_2. The reason for that is probably because of the limitation of CMRI, which can only provide 2D but not 3D imaging dataset.

Despite the superior endocardial definition, CMRI is still a 2D imaging modality, which scans the heart with discrete 2D slices of fixed slice thickness (8mm in this study). Between each two adjacent slices, there was no information for the FEM to adapt itself, thus it was impossible to judge whether the 3D approach of the LV endocardial surface was good or bad between slices, as shown in Figure 5-11(B_2I_2+FEM) and Figure 5-13 (B_2I_2+FEM). The main point of the FEM is that it can fully utilize the 3D geometric data together to approach the 3D surface. But if the data is not in 3D, the performance of the FEM is limited.

Although the FEM was limited in performing volume calculation, it still can provide reasonable endocardial border tracing. Figure 5-10 (B_2I_2+FEM) and Figure 5-12 (B_2I_2+FEM) showed that the delineation of the endocardial border by FEM is continuous and smooth.

5.5. Conclusion

For quantitative measurement of the human left ventricular volume through CMR images, the semi-automated border detection and interpolation algorithms were all feasible, easy to use, and were faster than using the manual method.

Without interpolation, the performance of border detection algorithm B_1 was better than B_2. It showed a higher correlation and smaller standard deviation in all the measurements, but both B_1 and B_2 significantly underestimated the left ventricular volume by a small amount compared with the gold standard.

Interpolation algorithm I_1 and I_2 were able to improve the quantification of left ventricular volumes and function. There were improvements for correlation in all the measurements, and also obvious improvements for bias and standard deviation in measuring EDV and ESV after using I_1 or I_2.

B_2+I_1 or I_2 showed superior performance than B_1+I_1 or I_2 did. The best combination of border detection and interpolation algorithm was B_2+I_2. B_2I_2 correlated best with the standard, and showed the smallest biases in measuring EDV and ESV.

The improvements via using I_1 and I_2 were not equally good in measuring EDV and ESV. After applying I_1 and I_2, there were still significant differences with the manual measurements in ESV for all algorithms, but not in EDV. This phenomenon caused the bias of SV and EF was bigger with interpolations than without interpolations.

Comparing with B_2I_2, the FEM did not provide superior performance in quantitative volume measurement of left ventricle through CMR images.

In conclusion, this study validated the semi-automated border detection and interpolation algorithms by comparing them to the manual method (gold standard) in quantification of LV volumes (EDV and ESV) and function (SV and EF) in human CMR images. The results proved that the border detection plus interpolation algorithms were able to provide comparable border delineation and volume calculation of LV cavities as the manual method. In the next chapter, these algorithms were applied on the clinical RT3DE images to access their performances.

Chapter 6. A preliminary study: quantitative LV volume measurements of clinical 3D echo images

In Chapter 3 and Chapter 4, the newly developed semi-automated surface tracing and volume calculation algorithms were applied to echo images of simple laboratory balloon phantoms and echo images of ultrasonically tissue-mimicking phantoms. The accuracy as well as the reproducibility of these algorithms was assessed by comparing the measured phantom volumes with known phantom volumes.

In Chapter 5, five different semi-automated interpolation algorithms were studied to delineate the LV endocardial borders. In order to validate them, both manual and semi-automated border delineations were applied to CMR images of 18 human subjects. The manually measured LV volumes were considered as gold standard, so that the accuracy of each interpolation algorithm was assessed.

The echo images have low spatial resolution, bad contrast and high noise, but the geometry of the phantoms is simple, symmetric and static. On the contrary, the geometry of human LV is complicated, asymmetric and dynamic, but the CMR images have high spatial resolution, good contrast and low noise. It had been proved that the algorithms did work well on those cases, although the performance of different algorithms might vary. However, now the question is: can the semi-automated algorithms work on low spatial resolution, bad contrast and high noise echo images of complicated human LV? In other words, are these algorithms applicable in clinical echo images?

In this chapter, the contrast enhanced stress echo tests were applied to five human subjects. 3D echo images of LV were acquired during the tests. Different semi-automated algorithms were applied to trace the LV endocardial surface and calculate the LV volumes. The performance of each technique was assessed and compared.

6.1. Imaging acquisition

Five patients with unknown cardiovascular disease participated in this study. Data were acquired during their clinical investigation by using a Philips iE33 Intelligent Echocardiography System. For each subject, the contrast enhanced real time 3D echo images were acquired in two stages during the test. The first acquisition was performed at the rest condition, and the second one was at the stress condition.

For each stage, two real time 3D images were acquired, which covered the human entire LV over one cardiac cycle (one R-R interval). Therefore, four real time 3D echo images were obtained for each subject, and in total 20 images were acquired. After that, the acquired images were transferred to a computer for off-line image analysis.

6.2. Image analysis

For each real time 3D image, the left ventricular endocardial surface was traced semi-automatically at the left ventricular end diastole and end systole. The LVEDV was measured after mitral valve closure, and the LV cavity reached its maximum volume. The LVESV was measured on the images with the smallest LV cavity. The EF was calculated as EF= (LVEDV-LVESV)/LVEDV.

The ROI initialization method 2 was used on 3D echo images. This is because the geometry of human LV is asymmetric. The ROI method 2 is more suitable than the method 1. Similar to the balloon phantom and tissue-mimicking phantom analysis, the 18 semi-automated algorithms (Table 3.1) were applied to the LV volume measurements. After that, the FEM was applied on the algorithm $G_2B_2I_2$.

6.2.1. Establish different ROIs

As introduced in Chapter 2, for all the newly developed semi-automated algorithms, the only manual intervention is the initialization of ROI. In ROI method 2, six points need to be selected manually (click by mouse), and three of them (first, third and fifth click) are in the centre area of LV blood pool, as shown in Figure 6-1 A.

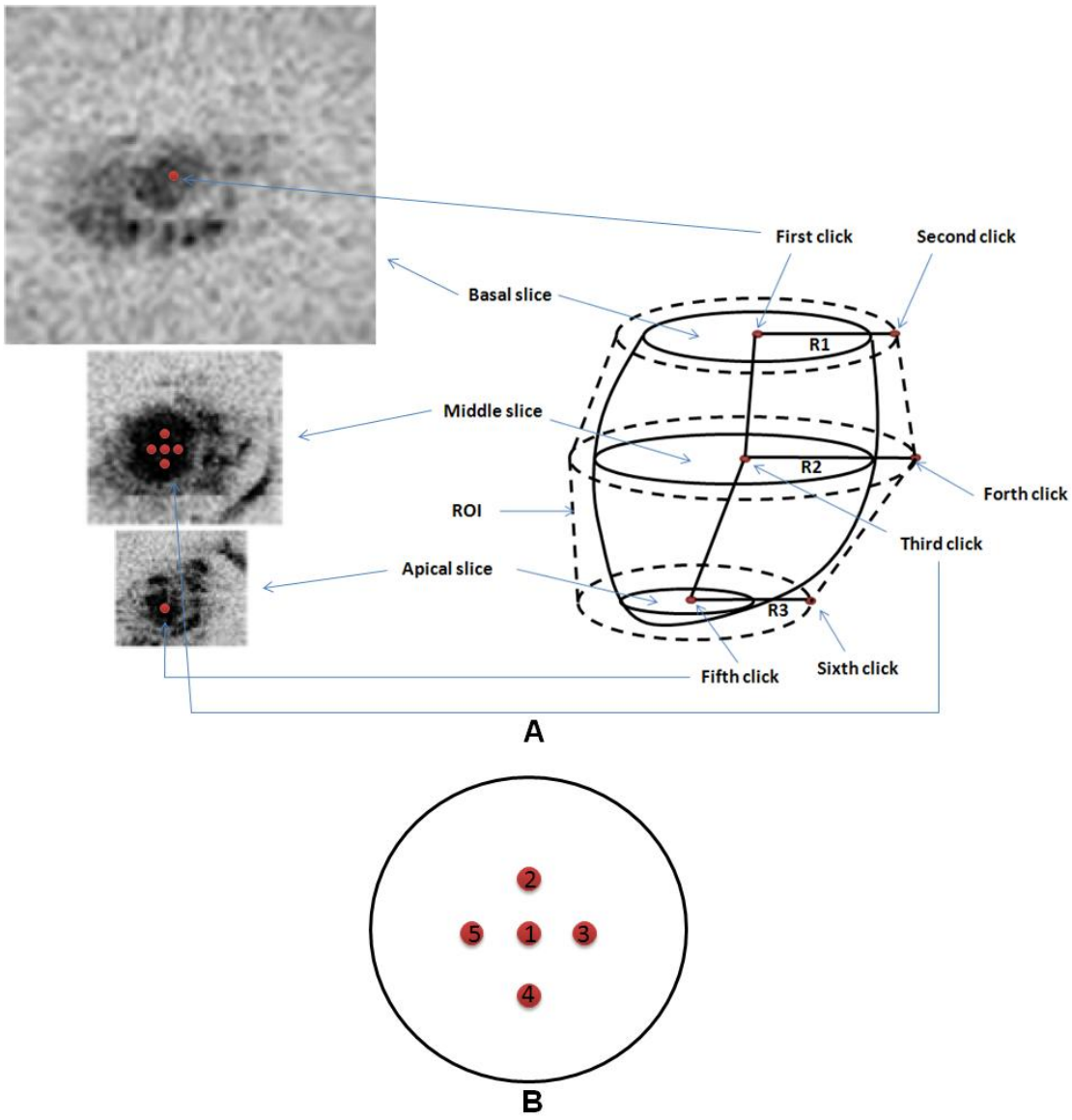


Figure 6-1 Establish five different ROIs by manually selecting five different points in middle slice. (A) A demonstration of how to build an ROI; (B) A sketch of the middle slice: To build five different ROIs, the manual selection of the point inside LV chamber follows this sequence: 1.centre-2.up-3.right-4.down-5.left.

In order to investigate the variability of these semi-automated algorithms, five different ROIs were established for every 3D image, so five LV volume measurements were performed for each algorithm. An ideal situation is that the five measurements are exactly the same, which means LV semi-automated measurement is independent of the ROI initialization.

The difference of the five ROIs was the third manual click in the middle slice (Figure 6-1 B). For each establishment, this point was selected deliberately in a certain sequence:

1.centre-2.up-3.right-4.down-5.left. Because of that, the central axis of each ROI varied five times, resulted in five different ROIs. However, no matter how the ROI changed, for each image, the LV was always contained inside the ROI.

6.2.2. Accuracy assessment

In the previous echo imaging analyses of balloon phantom, tissue-mimicking phantom and CMR imaging analysis of human subjects, the semi-automated volume measurements were compared with different references, such as known phantom volumes or the manually measured volumes. However, in this study there was no reference, so it was impossible to judge which algorithm was more or less accurate than others. But in all the previous analyses, one common finding was that the algorithms with interpolations showed better accuracy. Therefore, in this study, if significant difference was found between algorithms with and without interpolations, it was very likely that the algorithms with interpolations would be more accurate.

6.2.3. Reproducibility assessment

For each 3D echo image, five ROIs were established. Therefore, five repeated volume measurements were performed for every semi-automated algorithm. The variability of each algorithm was calculated as the coefficient of variation (CV, %): CV = SD/mean. The bigger CV indicated that the algorithm was less reproducible.

6.2.4. Comparison in rest and stress conditions

The real time 3D images were acquired during the clinical stress echo tests. For a subject, two images were firstly captured at the rest condition. After that, the dobutamine was injected to increase the subject's heart rate, so the heart reached stress condition. Then the other two were captured at this condition. For both conditions, the first acquired image was called Image 1, and the second one was called Image 2.

After previous assessments, algorithms with better accuracy and better reproducibility were found. Then one of the best algorithms was selected to investigate the difference of measuring the different images (Image 1 vs. Image 2), as well as the difference of measuring different clinical test conditions (rest vs. stress).

6.2.5. Statistical analysis

There were five human subjects and 18 new semi-automated algorithms in this study. Each subject was acquired four real time 3D echo images, two at rest and two at stress. Since each real time image covered one complete cardiac cycle (one R-R interval), LV volumes at end diastole and end systole were measured, and the EF was calculated. Because five ROIs were established for each volume measurement, in total there were 5 (subjects)×4 (real time images)×18 (algorithms)×5 (ROIs) = 1800 measurements obtained for measuring LVEDV, LVESV and EF respectively. ANOVA was used for an overall comparison of all the volume measurements. $P < 0.05$ was considered significant.

For comparison of the reproducibility, there were 1800/5=360 CV for measuring LVEDV, LVESV and EF respectively. ANOVA was used again to compare the performance of different algorithms. $P < 0.05$ was considered significant.

Paired Student t test and Bland Altman method were used to compare the difference of $G_2B_2I_2$ and $G_2B_2I_2$ with FEM in both volume measurements and reproducibility measurements. $P < 0.05$ was considered significant.

6.3. Results

Figure 6-2, Figure 6-3 and Figure 6-4 demonstrate some examples of LV endocardial surface delineation and 3D reconstruction done by four semi-automated algorithms at LVED and LVES. These pictures indicate that the algorithms can trace the LV endocardial surfaces successfully, although there were some differences of the tracing between different algorithms.

In order to compare the performance of different semi-automated algorithms, ANOVA was applied to all the measured volumes (1800 measurements and 360 CV for LVEDV, LVESV and EF respectively). The results are in the following sections.

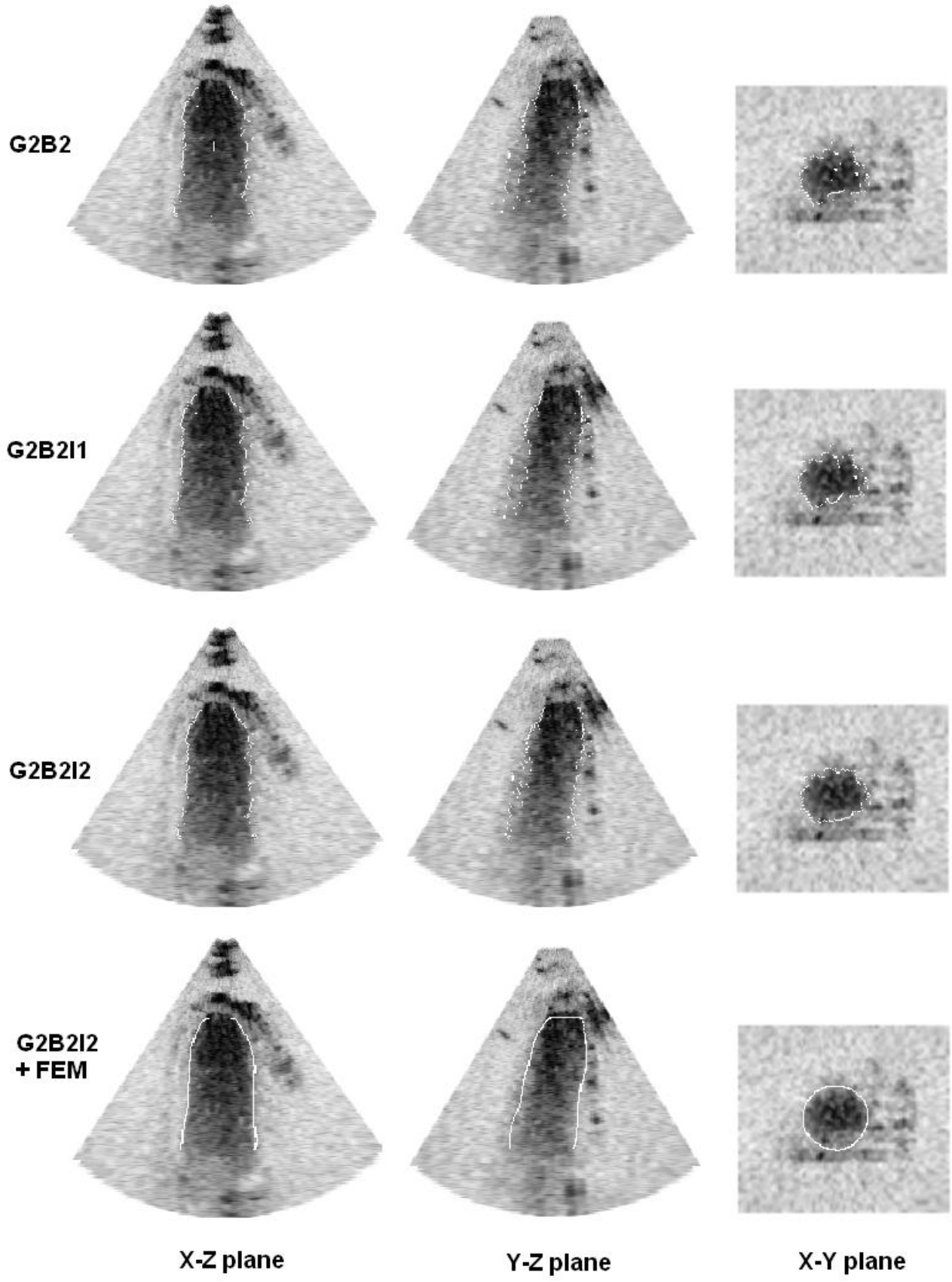


Figure 6-2 Three extracted planes showed surface delineation (white dots) at LVED by four algorithms (up to down: G_2B_2, $G_2B_2I_1$, $G_2B_2I_2$ and FEM)

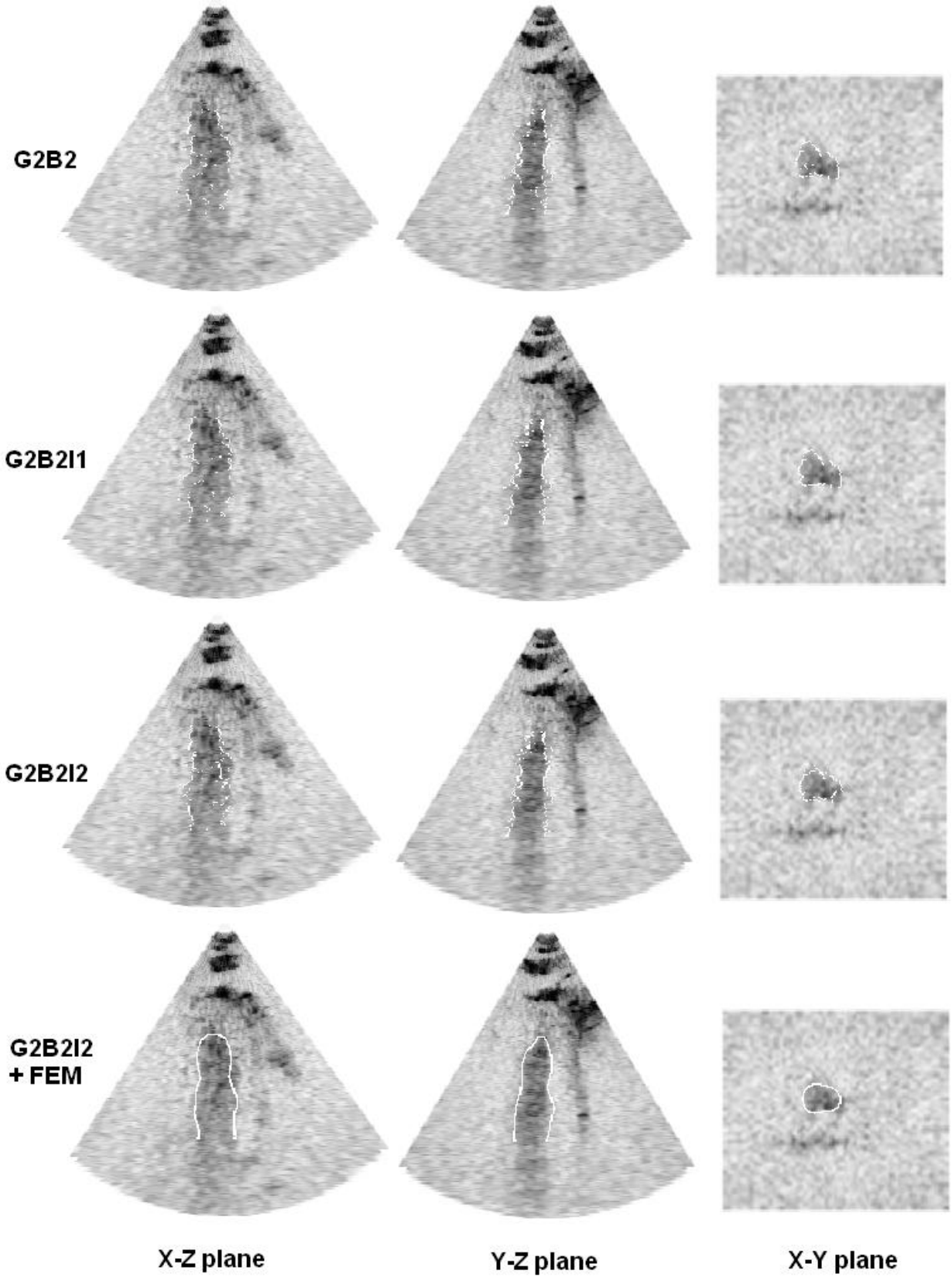


Figure 6-3 Three extracted planes showed surface delineation (white dots) at LVES by four algorithms (up to down: G_2B_2, $G_2B_2I_1$, $G_2B_2I_2$ and FEM).

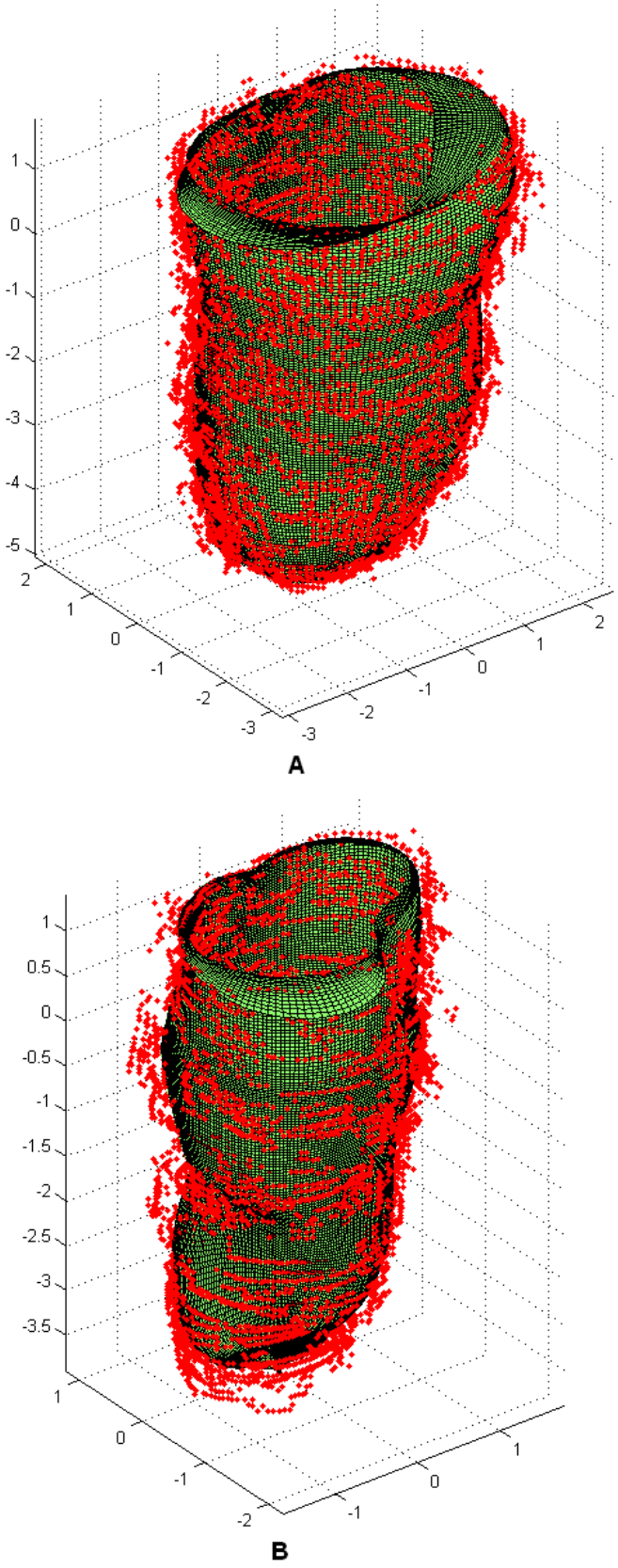


Figure 6-4 3D reconstruction of a LV endocardial surface at the LVED (A) and LVES (B) respectively: red dots: delineated by $G_2B_2I_2$; green meshes: delineated by using the FEM

6.3.1. Comparison of five ROIs

Figure 6-5 and Table 6-1 shows an overall comparison of the mean±95% CI of five repeated measurements given by five ROIs in LVEDV, LVESV and EF.

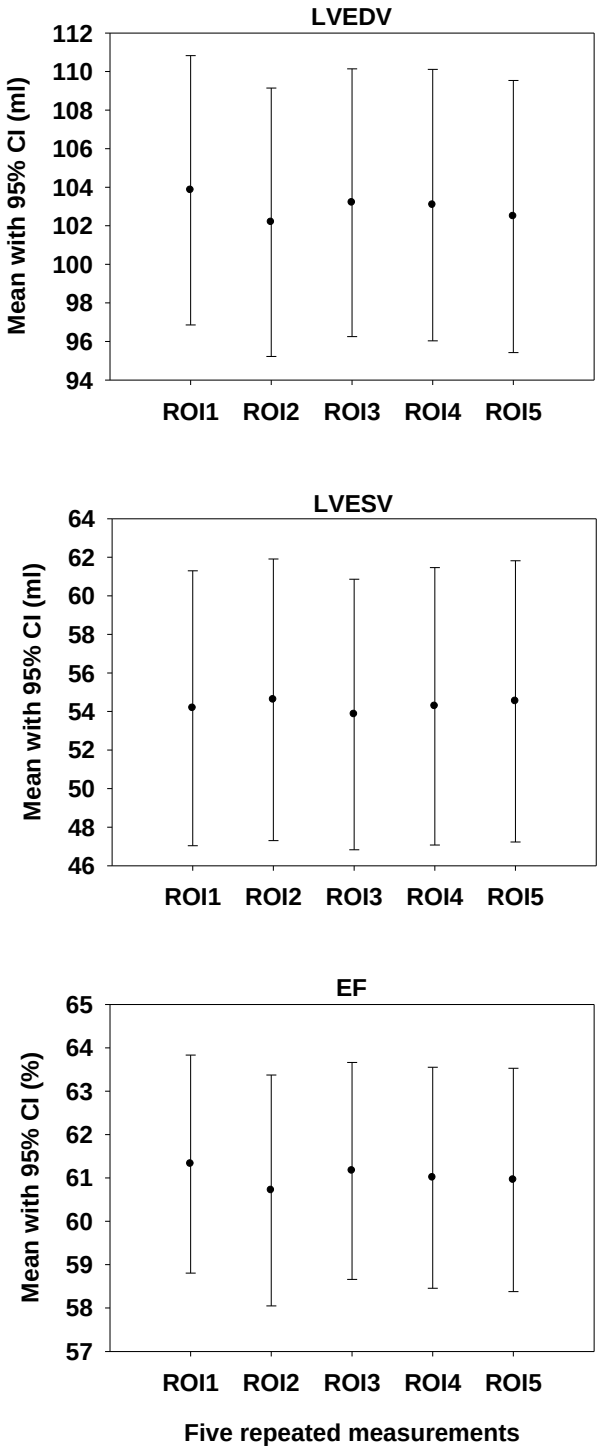


Figure 6-5 An overall comparison of mean±95% CI of repeated measurements given by five ROIs in LVEDV, LVESV and EF

All three charts in Figure 6-5 demonstrate that the five repeated measurements were similar. The error bars are almost overlapped to each other. The ANOVA indicated that there was no significant difference between any of the five measurements in LVEDV (test between subjects: P = 0.655), LVESV (test between subjects: P = 0.826) and EF (test between subjects: P = 0.738).

Table 6-1 Mean and 95% CI of repeated measurements given by five ROIs in LVEDV, LVESV and EF

LVEDV (ml)		
measurements	**mean**	**95%CI**
ROI1	103.8	±10.7
ROI2	102.2	±10.6
ROI3	103.2	±10.7
ROI4	103.1	±10.6
ROI5	102.5	±10.6
LVESV (ml)		
measurements	**mean**	**95%CI**
ROI1	54.2	±5.6
ROI2	54.6	±5.6
ROI3	53.8	±5.6
ROI4	54.3	±5.6
ROI5	54.5	±5.6
EF (%)		
measurements	**mean**	**95%CI**
ROI1	61.3	±6.3
ROI2	60.7	±6.3
ROI3	61.2	±6.3
ROI4	61.0	±6.3
ROI5	61.0	±6.3

6.3.2. Accuracy assessment

Figure 6-6 and Table 6-2 shows an overall comparison of the mean±95% CI of 18 semi-automated algorithms for measuring LVEDV, LVESV and EF. In Figure 6-6, it is quite clear that the performance of some algorithms were significantly different from the others, as some of the error bars are far away from the others. However, because of the lack of reference, it is difficult to say which algorithm was more accurate. But both Figure 6-6 and Table 6-2 show that volumes given by algorithms without interpolations

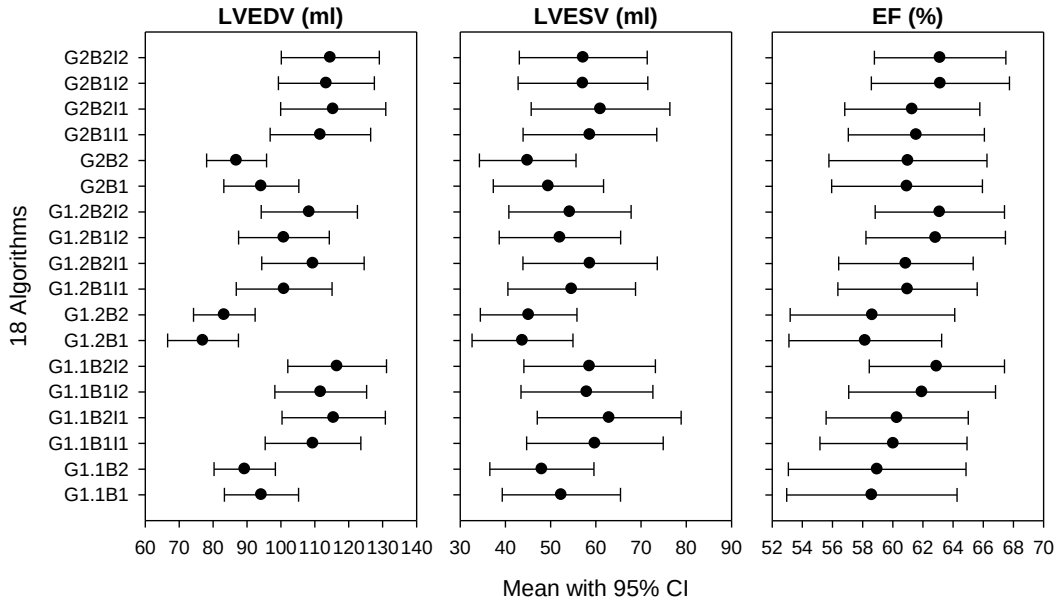


Figure 6-6 Mean±95% CI of 18 semi-automated algorithms for measuring LVEDV, LVESV and EF

Table 6-2 Mean±95% CI of 18 semi-automated algorithms for measuring LVEDV, LVESV and EF

	LVEDV (ml)		LVESV (ml)		EF (%)	
Algorithms	**Mean**	**95%CI**	**Mean**	**95%CI**	**Mean**	**95%CI**
$G_{1.1}B_1$	94.3	±18.5	52.4	±13.1	58.6	±5.6
$G_{1.1}B_2$	89.3	±17.5	48.1	±11.5	59.0	±5.9
$G_{1.1}B_1I_1$	109.5	±21.5	59.8	±15.1	60.0	±4.9
$G_{1.1}B_2I_1$	115.6	±22.6	63.0	±15.9	60.3	±4.7
$G_{1.1}B_1I_2$	111.7	±21.9	58.0	±14.6	61.9	±4.9
$G_{1.1}B_2I_2$	116.6	±22.8	58.6	±14.5	62.9	±4.5
$G_{1.2}B_1$	77.0	±15.1	43.8	±11.2	58.2	±5.1
$G_{1.2}B_2$	83.3	±16.3	45.1	±10.7	58.7	±5.5
$G_{1.2}B_1I_1$	101.0	±19.8	54.7	±14.1	61.0	±4.6
$G_{1.2}B_2I_1$	109.4	±21.5	58.7	±14.9	60.9	±4.5
$G_{1.2}B_1I_2$	100.9	±19.8	52.1	±13.4	62.8	±4.6
$G_{1.2}B_2I_2$	108.4	±21.2	54.3	±13.5	63.1	±4.3
G_2B_1	94.2	±18.5	49.5	±12.2	60.9	±5.0
G_2B_2	86.9	±17.0	44.9	±10.7	61.0	±5.2
$G_2B_1I_1$	111.6	±21.9	58.7	±14.8	61.6	±4.5
$G_2B_2I_1$	115.4	±22.6	61.0	±15.3	61.3	±4.5
$G_2B_1I_2$	113.4	±22.2	57.2	±14.3	63.2	±4.6
$G_2B_2I_2$	114.6	±22.5	57.2	±14.1	63.1	±4.4

($G_{1.1}B_1$, $G_{1.1}B_2$, $G_{1.2}B_1$, $G_{1.2}B_2$, G_2B_1 and G_2B_2) were much smaller than volumes given by the rest of algorithms.

Then, ANOVA was applied to measured LVEDV, LVESV and EF to further investigate the differences of the 18 algorithms. The results are as following:

For measuring LVEDV, significant difference was found between $G_{1.1}$ and $G_{1.2}$ (P=0.014; $G_{1.1}$, mean=106.2ml; $G_{1.2}$, mean=96.7ml), and also between $G_{1.2}$ and G_2 (P=0.015; G_2, mean=106.0ml). But $G_{1.1}$ was not significantly different from G_2 (P=0.974). No significant difference was found between two border detection algorithms B_1 and B_2 (P=0.360; B_1, mean=101.5ml; B_2, mean=104.4ml). For interpolations, there was significant difference between no interpolation and I_1 (P<0.001; No interpolation, mean=87.5ml; I_1, mean=110.4ml), and also between no interpolation and I_2 (P<0.001; I_2, mean=110.9ml). But I_1 was not significantly different from I_2 (P=0.895).

For measuring LVESV, no significant difference was found between any two of the three edge operators (test of between subjects: P=0.422, $G_{1.1}$, mean=56.6ml; $G_{1.2}$, mean=51.4ml; G_2, mean=54.8ml). No significant difference was found between two border detection algorithms B_1 and B_2 (P=0.867; B_1, mean=54.0ml; B_2, mean=54.6ml). For interpolations, there was significant difference between no interpolation and I_1 (P=0.003; No interpolation, mean=47.3ml; I_1, mean=59.3ml), and also between no interpolation and I_2 (P=0.026; I_2, mean=56.2ml). But I_1 was not significantly different from I_2 (P=0.441).

For measuring EF, the results were very similar as the results of measuring LVESV. No significant difference was found between any two of the three edge operators (test of between subjects: P=0.596, $G_{1.1}$, mean=60.5%; $G_{1.2}$, mean=60.8%; G_2, mean=61.9%). No significant difference was found between two border detection algorithms B_1 and B_2 (P=0.850; B_1, mean=60.9%; B_2, mean=61.1%). However, for the category of interpolations, there was no significant difference between no interpolation and I_1 (P=0.311; No interpolation, mean=59.4%; I_1, mean=60.8%), and nor between I_1 and I_2 (P=0.158; I_2, mean=62.9%). But no interpolation was significantly different from I_2

(P=0.015).

Based on the previous analysis of balloon phantoms and tissue-mimicking phantoms, algorithms without using interpolations (I_1 or I_2) were always less accurate than the others with interpolations. In this study, significant differences were found again between algorithms with and without using interpolations in measuring LVEDV and LVESV. Therefore, although there was no reference in this study, there is confidence that interpolation algorithms did improve the accuracy of measuring LV volumes significantly.

For algorithms with interpolations, since neither significant differences were found between two border detections (B_1 and B_2), nor between two interpolations (I_1 and I_2), the only category that made the differences was using of edge operators ($G_{1.1}$, $G_{1.2}$ and G_2) in measuring the LVEDV.

6.3.3. Reproducibility assessment

Since the lack of the reference of this study, it was difficult to judge the performance of each algorithm in the aspect of accuracy. Therefore, it was more important to investigate the reproducibility of each algorithm. Although in Section 6.3.1, the overall analysis of all 18 algorithms showed non-significant difference for five repeated measurements (five ROIs), the reproducibility of each algorithm can be calculated and compared by using the coefficient of variation (CV).

Figure 6-7 and Table 6-3 shows an overall comparison of the mean±95% CI of CV of 18 semi-automated algorithms for measuring LVEDV and LVESV. Figure 6-7 indicates that the CV is getting smaller and smaller from the bottom algorithm ($G_{1.1}B_1$) to the top algorithm ($G_2B_2I_2$).

After applying ANOVA, it was found that for measuring LVEDV, there were significant differences between any two of the three edge operators: $G_{1.1}$, mean of CV=5.1%; $G_{1.2}$, mean of CV=4.1%; G_2, mean of CV=2.6%; $G_{1.1}$ and $G_{1.2}$, P=0.001; $G_{1.1}$ and G_2, P<0.001; $G_{1.2}$ and G_2, P<0.001. Significant difference was also found between two border detection algorithms: P=0.001; B_1, mean of CV=4.4%; B_2, mean of CV=3.5%. There was no significant difference between two interpolation algorithms:

P=0.434; I_1, mean of CV=3.7%; I_2, mean of CV=3.5%. But algorithms without interpolations showed significant bigger CV: no interpolation, mean of CV=4.6%; no interpolation and I_1, P=0.005; no interpolation and I_2, P<0.001.

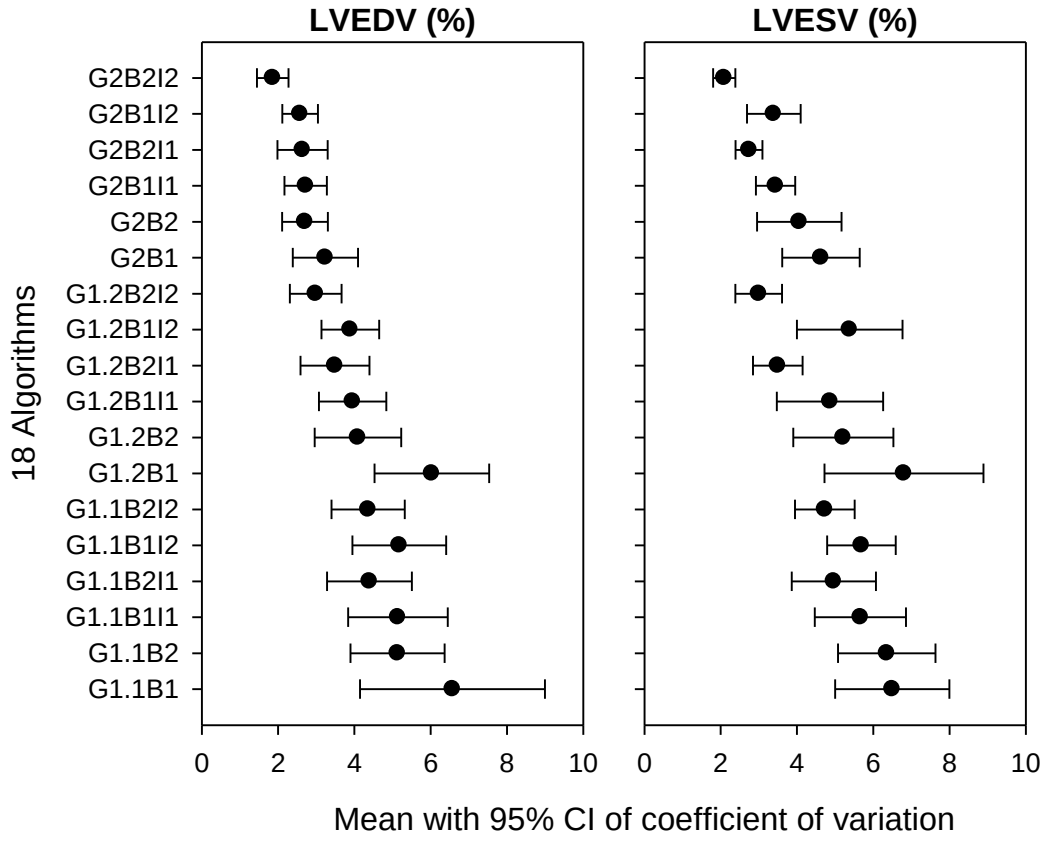


Figure 6-7 Mean±95% CI of coefficient of variation of 18 semi-automated algorithms for measuring LVEDV and LVESV

For measuring LVESV, the results were similar as the results of measuring LVEDV. Significant differences were found between any two of the three edge operators: $G_{1.1}$, mean of CV=5.7%; $G_{1.2}$, mean of CV=4.8%; G_2, mean of CV=3.4%; $G_{1.1}$ and $G_{1.2}$, P=0.008; $G_{1.1}$ and G_2, P<0.001; $G_{1.2}$ and G_2, P<0.001. Significant difference was also found between two border detection algorithms: P<0.001; B_1, mean of CV=5.2%; B_2, mean of CV=4.1%. There was no significant difference between two interpolation algorithms: P=0.646; I_1, mean of CV=4.2%; I_2, mean of CV=4.0%. But algorithms without interpolations showed significant bigger CV: no interpolation, mean of CV=5.6%; no interpolation and I_1, P<0.001; no interpolation and I_2, P<0.001.

Table 6-3 Mean±95% CI of coefficient of variation of 18 semi-automated algorithms for measuring LVEDV and LVESV

	CV for LVEDV (%)		CV for LVESV (%)	
Algorithms	**Mean**	**95%CI**	**Mean**	**95%CI**
$G_{1.1}B_1$	6.6	±2.9	6.5	±2.8
$G_{1.1}B_2$	5.1	±2.2	6.4	±2.8
$G_{1.1}B_1I_1$	5.1	±2.3	5.7	±2.5
$G_{1.1}B_2I_1$	4.4	±1.9	5.0	±2.2
$G_{1.1}B_1I_2$	5.2	±2.3	5.7	±2.5
$G_{1.1}B_2I_2$	4.4	±1.9	4.7	±2.1
$G_{1.2}B_1$	6.0	±2.6	6.8	±3.0
$G_{1.2}B_2$	4.1	±1.8	5.2	±2.3
$G_{1.2}B_1I_1$	4.0	±1.7	4.9	±2.1
$G_{1.2}B_2I_1$	3.5	±1.5	3.5	±1.5
$G_{1.2}B_1I_2$	3.9	±1.7	5.4	±2.4
$G_{1.2}B_2I_2$	3.0	±1.3	3.0	±1.3
G_2B_1	3.2	±1.4	4.6	±2.0
G_2B_2	2.7	±1.2	4.1	±1.8
$G_2B_1I_1$	2.7	±1.2	3.4	±1.5
$G_2B_2I_1$	2.6	±1.2	2.7	±1.2
$G_2B_1I_2$	2.6	±1.1	3.4	±1.5
$G_2B_2I_2$	1.9	±0.8	2.1	±0.9

From the above analysis, it is indicated that the 3D edge operator G_2 showed better reproducibility than the other two 2D operators ($G_{1.1}$ and $G_{1.2}$). The border detection B_2 showed better reproducibility than B_1. Compared with no interpolation algorithms, I_1 and I_2 did not only improve the accuracy, but also improved the reproducibility of LV volume measurements. Therefore, algorithm $G_2B_2I_1$ and $G_2B_2I_2$ were probably the best two among all the 18 semi-automated algorithms. In the following investigation, $G_2B_2I_2$ was used to investigate the difference of measuring LV volumes of repeated images (Image 1 vs. Image 2) and different clinical conditions (rest vs. stress).

6.3.4. Comparison in rest and stress conditions

For algorithm $G_2B_2I_2$, there were 100 measurements of LVEDV, LVESV and EF respectively (n=5 subjects×4 images×5 ROIs). Table 6-4 shows the heart rate of the five subjects when the 3D echo images were acquired. The heart rates were almost the same

when the two images were captured at the same condition, but rose significantly when the condition changed from rest to stress. ANOVA was applied to compare the difference of measured LV volumes by $G_2B_2I_2$.

Table 6-4 Subjects heart rate when the 3D echo images were acquired

Subjects	Images	Condition	Heart Rate (bpm)
1	Image 1	Rest	82
	Image 2	Rest	80
	Image 1	Stress	126
	Image 2	Stress	124
2	Image 1	Rest	55
	Image 2	Rest	53
	Image 1	Stress	79
	Image 2	Stress	71
3	Image 1	Rest	63
	Image 2	Rest	62
	Image 1	Stress	82
	Image 2	Stress	79
4	Image 1	Rest	89
	Image 2	Rest	91
	Image 1	Stress	113
	Image 2	Stress	113
5	Image 1	Rest	86
	Image 2	Rest	90
	Image 1	Stress	121
	Image 2	Stress	110

For measuring LVEDV, the five repeated measurements of five ROIs showed non-significant difference (test between repeated measurements: P=0.999; mean: ROI 1=114.7ml, ROI 2=114.7ml, ROI 3=114.9ml, ROI 4=114.2ml and ROI 5=114.3ml; standard error of the mean difference=3.0). This was consistent with the finding of overall ANOVA of the 18 algorithms in Section 6.3.1. The measurements of two repeated images (Image 1 and Image 2) at the same condition did not show significant difference either (test between repeated images: P=0.760; mean: Image 1=114.3ml, Image 2=114.9ml; standard error of the mean difference=1.9). However, the measured LVEDV significantly decreased from rest to stress (test between conditions: P=0.008; mean: rest=117.1ml, stress=112.0ml; standard error of the mean difference=1.9ml).

For measuring LVESV, the five repeated measurements of five ROIs showed non-significant difference (test between repeated measurements: P=0.892; mean: ROI 1=56.6ml, ROI 2=57.6ml, ROI 3=57.2ml, ROI 4=56.9ml and ROI 5=57.8ml; standard error of the mean difference=1.3). The measurements of two repeated images (Image 1 and Image 2) at the same condition did not show significant difference either (test between repeated images: P=0.764; mean: Image 1=57.1ml, Image 2=57.4ml; standard error of the mean difference=0.8). Similar to the finding of measuring LVEDV, the measured LVESV also significantly decreased from rest to stress (test between conditions: P<0.001; mean: rest=60.1ml, stress=54.4ml; standard error of the mean difference=0.8ml).

For measuring EF, the five repeated measurements of five ROIs showed non-significant difference (test between repeated measurements: P=0.990; mean: ROI 1=63.5%, ROI 2=63.0%, ROI 3=63.1%, ROI 4=63.2% and ROI 5=63.0%; standard error of the mean difference=1.2). The measurements of two repeated images (Image 1 and Image 2) at the same condition did not show significant difference either (test between repeated images: P=0.807; mean: Image 1=63.0%, Image 2=63.2%; standard error of the mean difference=0.8). However, different from the findings of measuring LVEDV and LVESV, the measured EF significantly increased from rest to stress (test between conditions: P<0.001; mean: rest=61.3%, stress=65.0%; standard error of the mean difference=0.8).

6.3.5. Compare $G_2B_2I_2$ and FEM

Figure 6-8 showed the comparison between $G_2B_2I_2$ and $G_2B_2I_2$ with FEM in measuring LVEDV, LVESV and EF respectively by Bland Altman method. It should be noted that the five subjects are patients attending for clinical investigation. Therefore, they might have abnormal cardiac parameters.

There were no significant differences in measuring LVEDV (P=0.057; mean difference (95%CI): -0.4(±0.4)ml) and EF (P=0.909; mean difference (95%CI): -0.0(±0.1)%). However, they are significantly different in measuring LVESV (P=0.014; mean difference (95%CI): -0.4(±0.3)ml).

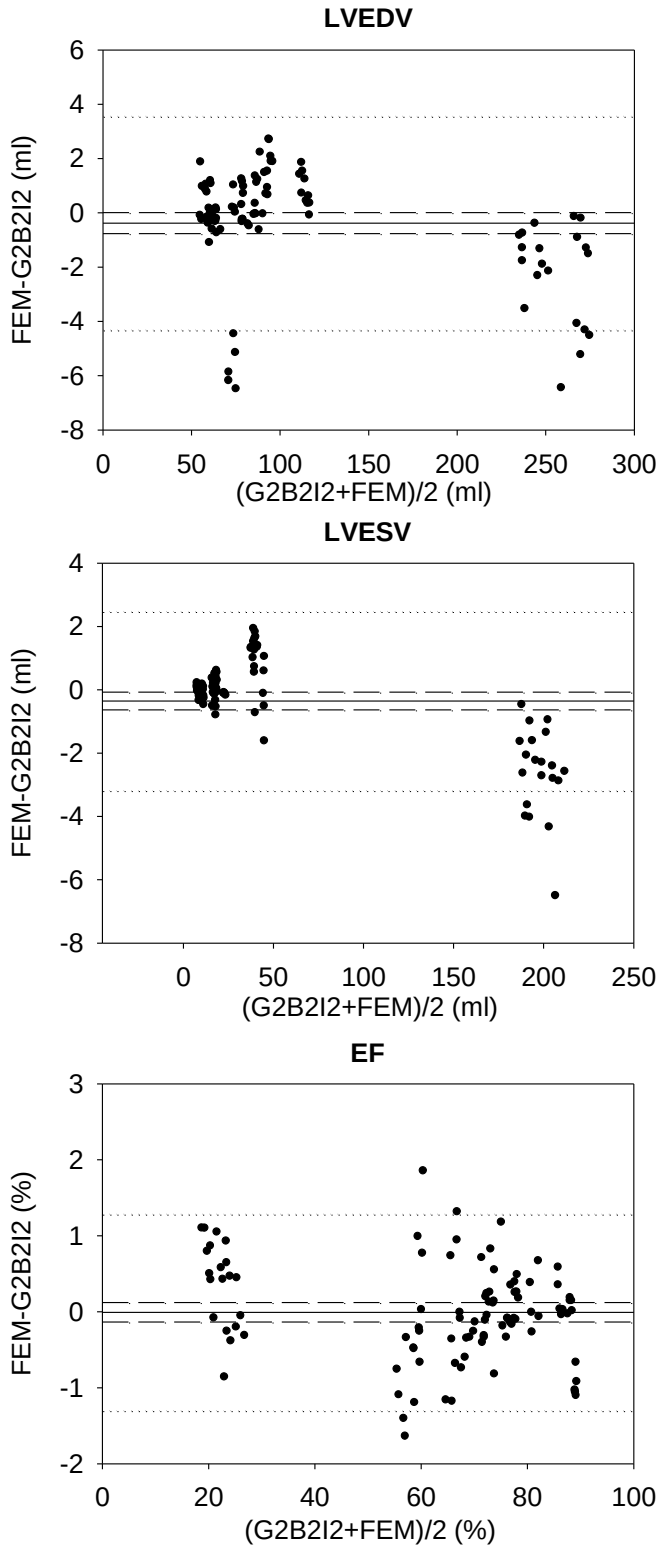


Figure 6-8 Bland Altman plots: comparison between $G_2B_2I_2$ and $G_2B_2I_2$ with FEM in measuring LVEDV, LVESV and EF respectively. Solid lines: bias (mean difference); dash lines: 95% confidence interval (CI) of bias; dotted lines: upper and lower limits of agreement

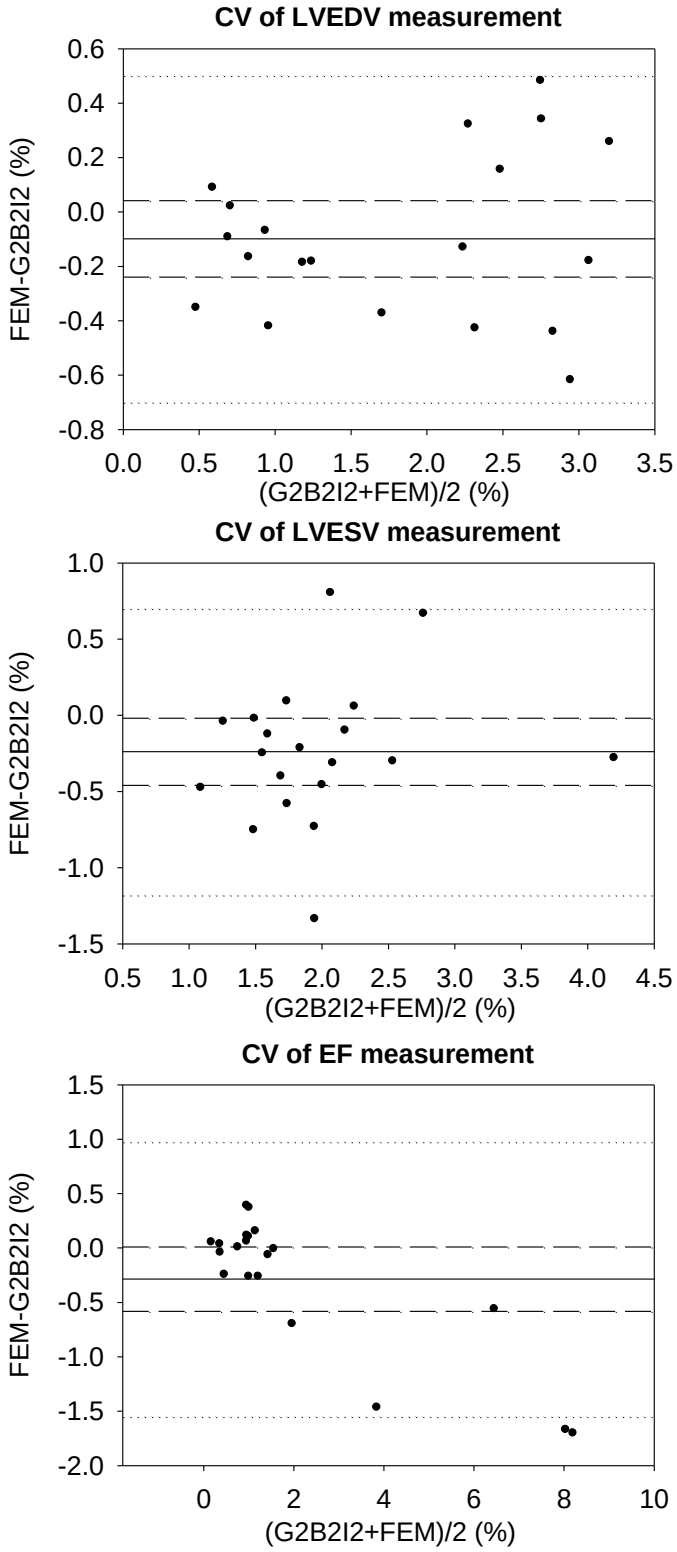


Figure 6-9 Bland Altman plots: comparison of coefficient of variation (CV) between $G_2B_2I_2$ and $G_2B_2I_2$ with FEM in measuring LVEDV, LVESV and EF respectively. Solid lines: bias (mean difference); dash lines: 95% confidence interval (CI) of bias; dotted lines: upper and lower limits of agreement

The results of reproducibility comparison were very similar with the results of volume measurement comparison (Figure 6-9). The coefficient of variations (CV) of two algorithms showed no significant differences in measuring LVEDV (P=0.157; mean difference (95%CI): -0.09(±0.13)%) and EF (P=0.056; mean difference (95%CI): -0.28(±0.28)%). But significant difference of CV was found in measuring LVESV (P=0.034; mean difference (95%CI): -0.24(±0.20)%).

The only significant difference between $G_2B_2I_2$ with and without FEM was in measuring LVESV. Because of the lack of reference, it was impossible to judge which algorithm was more accurate. However, after using the FEM, the CV of all measurements decreased a little, as shown in Figure 6-9. This indicated that FEM further improved the reproducibility.

6.4. Discussion

In the previous studies, the newly developed semi-automated algorithms were tested and compared by measuring volumes of balloon phantoms and tissue-mimicking phantoms from 3D echo images. However, comparing with real human hearts, the geometry of the phantoms is simple and symmetric. The human heart is moving all the time, which makes the imaging analysis much more difficult. It was still a question whether the algorithms were capable of delineating the endocardial surface of left ventricles from real clinical 3D echo images of human subjects. One of the aims of this study was to answer that question.

Figure 6-2 and Figure 6-3 show some examples of the semi-automated delineations of a LV endocardial surface from a contrast enhanced 3D echo image. It is very clear that the four algorithms showed in the figures succeeded in detecting most of the muscle wall boundaries. After the initialization of the ROIs, all the algorithms including the $G_2B_2I_2$ with FEM can delineate the LV endocardial surface automatically in 3D echo images. However, because of the influence of noise and artefact in echo images, the complexity of LV geometry, the roughness of the LV surface (especially at the LVES), the boundary delineation of every algorithm distinguished each other in many details,

such as the examples in Figure 6-2 and Figure 6-3. Nevertheless, this study proved that all the semi-automated algorithms were applicable for LV endocardial surface delineations and LV volume measurements in clinical 3D echo images.

Similar to the previous studies, the performance of the semi-automated algorithms was compared in the respect of accuracy and reproducibility. Some of the findings were the same as the previous studies, but some were different. This is discussed in the following sections.

6.4.1. Difference of three edge detection operators

One of the drawbacks of this study was the lack of a reference. Therefore, it was difficult to say which algorithm was more or less accurate than the others. But, based on the previous studies of balloon and tissue-mimicking phantoms, it was found that the performance of edge detection operators was related with the operator size and dimensions.

The volume measurements of $G_{1.1}$ (5×5) and G_2 ($5\times5\times5$) were significantly larger than the measurements of $G_{1.2}$ (11×11). By comparing with the known volumes of the phantoms, $G_{1.2}$ showed a bigger bias (significant underestimation) than the other two operators. It was probably the edge broaden effect from the bigger operator size that caused such volume underestimation. The same phenomenon was found in this study for measuring LVEDV: the measurements of $G_{1.2}$ were significantly smaller than the other two operators (section 6.3.2).

Referring to the comparison of reproducibility, the bigger operator size showed certain advantages. Since a central pixel can utilise more information from its neighbour because of a bigger operator size, the influence from imaging noise was reduced. This made the volume measurements more repeatable. In the study of balloon phantoms, tissue-mimicking phantoms as well as this study of clinical images, the reproducibility of $G_{1.1}$ was significantly worse than the other two, and G_2 showed the best reproducibility among the three.

For the comparison of edge operators in clinical 3D echo images, there was no doubt that the 3D edge operator can provide better reproducibility than the conventional 2D

edge operators. In the aspect of accuracy, based on the previous phantom studies, it was likely that the 3D operator was also better than the 2D operators. Therefore, to make full use of 3D echo images, the imaging analysis should be performed in all three dimensions, rather than only analyse consecutive 2D slices. This is one of the most important findings in this series of 3D echo imaging studies.

6.4.2. Difference of two border detection algorithms

The different performances of two border detection algorithms (B_1 and B_2) in 3D echo images were detected in all three studies (Table 6-5). In the balloon phantom study, B_1 showed significantly smaller bias with known volumes, and slightly better reproducibility than B_2 did. In the tissue-mimicking phantom study, however, B_2 showed better accuracy, but no significant difference of reproducibility was detected between B_1 and B_2. In this study, B_1 was not significantly different from B_2 in LV volume measurements, but B_2 was significantly repeatable than B_1.

Table 6-5 Different performances of two border detection algorithms (B_1 and B_2) in 3D echo images of three studies

Studies	Accuracy	Reproducibility
Balloon phantoms	B_1 was better.	B_1 was better.
Tissue-mimicking phantoms	B_2 was better.	No significant difference
Clinical subjects	No significant difference	B_2 was better.

Three different studies showed three different conclusions. Therefore, it is difficult to judge which border detection algorithm was better. Theoretically, if the imaging quality is good: the definition of the wall boundary is sharp and clear, and the imaging noise and artefact is low, B_1 can provide better performance. On the contrary, if the definition of the wall boundary is fuzzy, and significant noise and artefact appears, B_2 can provide better performance. It was probably because two different echo imaging systems were used in the three studies (Philips Sonos 7500 in study 1 and Philips iE33 in study 2 and 3), and the imaging subjects were in different environments (balloon phantoms in water, tissues-mimicking phantoms in saline water and hearts in human chests), that resulted in

different imaging qualities.

But the purpose of developing the border detection algorithms was to detect the LV endocardial surface boundaries of human subjects in 3D echo images. Therefore, further study of more clinical echo images of human subjects needs to be carried out to investigate the differences of the two border detection algorithms.

6.4.3. The interpolation algorithms

Surprisingly, the results showed no significant difference between the two interpolation algorithms I_1 and I_2 in all three studies, although the border delineation of the two algorithms was obviously different. One example is showed in Figure 6-2 and Figure 6-3. The linear and elliptical interpolation algorithms were designed with different geometric principles, as introduced in section 2.4. Therefore, although the edge detection operators and border detection algorithms were the same in the two figures, the pictures of $G_2B_2I_1$ and $G_2B_2I_2$ (pictures in row 2 and 3) demonstrate different border delineations. However, the results indicated that the total number of voxels inside the LV counted by the two surface delineations were almost the same.

Despite no difference detected between the two interpolations, there was indeed a big difference between using and not using interpolations in all three studies. Based on the balloon phantom and tissue-mimicking study, the accuracy of volume measurements were significantly improved by applying interpolation algorithms. It was partially because the border detection algorithms did not detect any borders in some artefact area, where the wall boundaries simply disappeared. It was also partially because the border detection algorithms detected some voxels with high noise, and considered them as wall boundaries. The same phenomenon happened in this study too. The algorithms with interpolations showed much larger volume measurements than those without interpolations.

One of the biggest problems of 3D echo images is the high noise and artefact, which makes the semi-automated or automated LV border delineation very challenging. Border detection algorithms B_1 and B_2 tried to delineate boundaries purely based on the imaging intensity differences between the two sides of the boundaries. However,

because of the noise and artefact, such efforts did not work effectively in almost all 3D echo images. There must be some extra information besides the imaging intensity changes guiding the boundary delineation, which is exactly the function of interpolation algorithms. They use the geometric assumptions of human LV morphology to provide extra information for boundary delineations. In this study, it proved that both I_1 and I_2 helped improving the LV surface tracing in clinical 3D echo images.

6.4.4. The finite element model

Figure 6-2, Figure 6-3, and Figure 6-4 showed examples of LV endocardial surface delineation given by $G_2B_2I_2$+FEM. Comparing with delineations from $G_2B_2I_2$, FEM provided smoother LV surfaces in all three dimensions.

The continuity of a LV surface delineation between adjacent slices along its long axis was not clear by border detection and interpolation algorithms, as shown in the first three rows of Figure 6-2 and Figure 6-3. This was because the boundary tracing was performed separately, slice after slice. The FEM utilised the surface delineation by $G_2B_2I_2$ from all three dimensions to insure the surface continuity. The idea behind it was very similar with 3D edge operator, which utilised the voxel information from all three dimensions in a 3D echo image to reduce the influence of noise and artefact.

In the balloon phantom and tissue mimicking phantom studies, the results showed significant improvement of the accuracy after applying the FEM. In this study, significantly difference of volume measurements between using and not using FEM was only found in measuring LVESV. However, because of the lack of reference, it was difficult to judge which one was more accurate. But using FEM also showed significant better reproducibility in measuring LVESV. To determine whether there is improvement of using the FEM, further study of clinical 3D echo images with a reliable reference was needed.

6.5. Conclusion

The semi-automated algorithms can delineate LV endocardial surfaces and measure LV volumes and EF quantitatively in clinical 3D echo images.

The difference of manual initialization of the ROIs did not influence the performance of all the semi-automated algorithms significantly.

Significant difference was found between 2D and 3D edge operators. But because of the lack of reference, the accuracy of each edge operator cannot be evaluated. However, the 3D edge operator did provide better reproducibility in repeated measurements compared with the conventional 2D edge operators.

The border detection algorithm B_1 and B_2 did not show significant difference in LV volume measurements. However, the reproducibility of B_1 was better than B_2 in this study.

Border interpolations were necessary. After applying interpolation algorithms, the LV surface delineation was improved. But for the accuracy and reproducibility of measuring LV volumes, the performance of the two interpolation algorithms (I_1 and I_2) did not show any significant difference.

For measuring LVESV, there was significant difference between $G_2B_2I_2$ and FEM. FEM improved the reproducibility. But for measuring LVEDV and EF, both algorithms provide the same performance.

The semi-automated algorithm $G_2B_2I_2$ indicated that for the clinical subjects, there was no significant difference between the repeated measurements at the same condition (rest or stress). However, from the rest to stress, both LVEDV and LVESV decreased, which resulted in an increase of EF.

Chapter 7. Conclusions and future work

There were two aims of this project. Firstly, it aimed to develop semi-automated algorithms for delineation of LV endocardial surfaces and quantification of LV volumes and EF from clinical real time 3D echo images. Secondly, it aimed to investigate what factors and how these factors in 3D echo images influence the quantitative LV volume measurements, by assessing and comparing the performance of these algorithms.

To reach these two aims, four different studies were designed and carried out: the laboratory balloon phantom study, the tissue-mimicking study, the CMRI validation study and the clinical stress echo study. This chapter summarises all the studies, gives final conclusions and discusses future work.

7.1. Summary

Different semi-automated algorithms were developed, and different studies were designed to assess their performance. This section gives a brief summary of the algorithms and the studies.

7.1.1. Summary of the semi-automated algorithms

Different semi-automated algorithms were designed and developed. They all following the same procedure: establishing the region of interests (ROI), applying the edge detection operator, applying the border detection algorithm, using (or not using) the border interpolation algorithm, and finally, applying the finite element model (FEM).

Two different methods were designed for establishing the ROI. In order to minimize the manual intervention, only two or six points were manually selected for the two methods respectively. For each study, only one of the methods was applied.

There were three edge operators with different sizes and dimensions: $G_{1.1}$ (5×5), $G_{1.2}$ (11×11) and G_2 ($5\times5\times5$), two border detection algorithms: B_1 and B_2, and two interpolation algorithms: I_1 and I_2. Together, they formed 18 different semi-automated algorithms (Table 3-1).

The FEM is a mathematical model, which was specifically designed for heart ventricle geometry by using 3D finite element meshes referred to a prolate spheroid coordinate system. It was not applied for all the 18 algorithms, but only to the algorithm (algorithms) that showed better performance in volume measurements.

After the delineation of chamber wall surfaces, the chamber volumes were calculated easily by the summation of discs method.

7.1.2. Summary of the four studies

In order to assess and compare the accuracy and reproducibility of each algorithm, different studies were designed and carried out. The first study was the balloon phantom study. A balloon filled with water of different volumes was imaged by a 3D echo system. The semi-automated algorithms were applied to measure the balloon volumes from 3D

echo images. To assess the accuracy, the measured volumes were compared with known volumes. To assess the reproducibility, each algorithm performed five repeated measurements on five 3D echo images of each phantom.

The results proved that the semi-automated algorithms were able to trace the balloon inner wall surfaces and measure volumes on 3D echo images. The manual initialization was very simple, and the algorithms were easy to apply. Different algorithms showed different performance in the measurements. However, all the algorithms showed significant underestimation of the known phantom volumes. There were some possible reasons for that. One of them was the difference of the ultrasound propagation speed in different materials. To overcome this limitation, tissue-mimicking phantoms were developed and used.

The tissue-mimicking phantoms used a special-made material that provided the same ultrasound property as human blood and soft tissue. The size and geometry of each phantom was precisely controlled by a CNC (computer numerical control) lathe system.

A conventional 2D echo system and a later 3D echo system were used for imaging the phantoms separately. The semi-automated algorithms were applied to both 2D and 3D images to delineate the phantom inner wall surfaces and measure phantom volumes. To assess the accuracy, the measured volumes were compared with known volumes. To assess the reproducibility, each algorithm performed two repeated measurements on two images (2D or 3D images) of each phantom.

The results showed that the measurements on 3D echo images provided superior accuracy than conventional 2D images. The performance of each algorithm was quite similar with the balloon phantom study, except for border detection algorithms.

Besides the studies on 3D echo images, the border detection and interpolation algorithms were validated by applying them on the clinical CMR (cardiac magnetic resonance) images to trace human LV endocardial surfaces and measure LV volumes (LVEDV and LVESV) and function (SV and EF). The results were compared with the gold standard manual tracing measurements.

The results proved that for quantitative measurement of LV volumes and function

through CMR images, the semi-automated border detection and interpolation algorithms were easy to use and faster than using the manual method. They were also able to provide comparable border delineation and volume calculation of LV cavities as the manual method.

The last study of this project was the measurement of LV volumes and function on contrast enhanced 3D echo images from clinical stress echo test. Images were collected from five subjects, and their LVEDV, LVESV and EF were measured by the semi-automated algorithms. Since there was no reference to compare with, the main index to assess the algorithm performance was the reproducibility.

For each subjects, two real time 3D echo images were acquired during the rest condition, and another two during the stress condition. For each real time image, every semi-automated algorithm measured the LVEDV and LVESV five times from five different ROIs.

The results proved again that the semi-automated algorithms can delineate LV endocardial surfaces and measure LV volumes and EF quantitatively in clinical 3D echo images.

The difference of manual initialization of the ROIs did not influence the performance of all the semi-automated algorithms significantly. There was no significant difference between the repeated measurements at the same condition (rest or stress). However, from the rest to stress, both LVEDV and LVESV decreased, which resulted in an increase of EF.

7.2. Conclusions

The newly developed semi-automated algorithms can delineate the human LV endocardial surfaces, measure LV volumes and EF on clinical real time 3D echo images. The only manual intervention was the establishment of ROI, which was designed to be simple and easy to perform.

The 3D edge operator G_2 provided the best accuracy as well as the best

reproducibility among the three edge operators. For the 2D operators, $G_{1.1}$ showed better accuracy, but $G_{1.2}$ showed better reproducibility.

The border detection algorithms B_1 and B_2 offered different performances in different studies. Generally, without any interpolations, B_1 showed better accuracy. However, after using interpolation algorithms, B_2 showed better accuracy and/or reproducibility. Therefore, B_2 was considered to be better at cooperating with interpolation algorithms.

The interpolation algorithms I_1 and I_2 improved the accuracy significantly. However, surprisingly, both algorithms showed similar performance in terms of both accuracy and reproducibility. No significant difference was found between them in all the studies.

In the balloon phantom study and tissue-mimicking study, after applying the FEM, the accuracy was improved. In the clinical stress echo study, although significant difference was found between using and not using FEM in measuring LVESV, because of the lack of reference, it was impossible to judge whether the accuracy was improved or not. However, the reproducibility of measuring LVESV was improved. In the CMRI validation study, there was no significant difference between using and not using FEM. This was probably because the CMR images were discrete 2D images.

To sum up all the studies, it was found that: In order to achieve a better quantitative measurement of 3D echocardiography images, the imaging analysis should be performed from all three dimensions, as this utilized the information contained in 3D images, rather than split them into multiple 2D slices and analyzed them separately. Analyzing 3D echo images in 3D can minimize the negative influence of noise and artefact, and so improve the accuracy as well as reproducibility.

The noise and artefact influenced the semi-automated delineation of phantom inner wall surfaces and human LV endocardial surfaces. Besides that, the imaging spatial resolution and edge definition also played important roles in the aspect of accuracy. Even if the borders were delineated correctly on images, because the poor spatial resolution caused the edge spread effect, volumes were still underestimated. Therefore, the border delineation was suggested to extend towards outside for certain distance from the original detected borders. In the balloon phantom study, it was extended by three

pixels (voxels). In the tissue-mimicking phantom and clinical stress echo study, it was extended by one pixel (voxel). In the CMRI study, because MRI provided superior LV endocardial boundary definition, there was no extra extension applied.

7.3. Future work

In the clinical stress echo study, data were collected only from five subjects. In the future, data will be collected from more subjects. And because there was no reference in this study, it was impossible to know how accurate the semi-automated measurements can be. One possible solution is to apply the clinical bi-plane measurements on the same group of images, and compare the semi-automated measurements with clinical manual measurements. Other solution is using the commercial available software, such as QLAB, to perform the measurements, and compare the semi-automated algorithms with the commercial software package.

The 3D echocardiography can provide real time 3D echo images covering a completed cardiac cycle. However, the semi-automated algorithms can only measure one volume of one frame, after the ROI was set up. In the future, it will be interesting to further develop the algorithms, so that after the establishment of two ROIs at LVED and LVES, the algorithms can perform a series of volume measurements on every frame of one completed cardiac cycle automatically. The LV volume variation can be traced during one heart beat.

To make full use of the excellent temporal resolution of the real time 3D echocardiography, the semi-automated algorithms can be developed to perform 4D analysis, which is measuring LVEDV and LVESV though more than two cardiac cycles of one subject. Because the noise and artifact appearing in one heart beat may not appear in the following beats, the average of the measurements over several cardiac cycles should be more accurate and reliable.

After developing algorithms for tracing LV endocardial surface, new algorithms can be designed to trace the LV epicardial surface, calculate LV mass and trace LV myocardial regional and global movements.

References

1. **Kisslo JA, Adams DB, Leech GJ.** Two-dimensional echocardiography in the normal heart. *Duke University Medical School Echo in Context Teleconferences* 2000.
2. **Reef V, Horst J.** Echocardiography Tutorial. *University of Pennsylvania, School of Veterinary Medicine* 1999
3. **Corsi C, Borsari M, Consegnati F, Sarti A, Lamberti C, Travaglini A, Shiota T, Thomas JD.** Left ventricular endocardial surface detection based on real-time 3D echocardiographic data. *Eur J Ultrasound* 2001; 13: 41-51.
4. **Suetens P.** Fundamentals of medical imaging, first edition. *Cambridge University Press* 2002.
5. **Krishnammorthy VK, Sengupta PP, Gentile F, Khandheria BK.** History of echocardiography and its future applications in medicine. *Crit Care Med* 2007; 35: S309-13.
6. **"http://www.cardiovascularultrasound.com/content/supplementary/1476-7120-5-47-S2-posterframe.jpg"** available on 25th, Oct 2008.
7. **Lang RM, Mor-Avi V, Sugeng L, Nieman PS, Sahn DJ.** Three-dimensional echocardiography: the benefits of the additional dimension. *J Am Coll Cardiol* 2006; 48: 2053–69.
8. **Mor-Avi V, Lang MR.** The use of real-time three-dimensional echocardiography for the quantification of left ventricular volumes and function. *Curr Opin Cardiol* 2009; 24(5): 402-9.
9. **Rogers WJ, Canto JG, Lambrew CT**. Temporal trends in the treatment of over 1.5 million patients with myocardial infarction in the U.S. from 1990 through 1999: the national registry of myocardial infarction 1, 2 and 3. *J AM Coll Cardiol* 2000; 36: 2056-63.
10. **John Sutton MS, Pfeffer MA, Plappert T, Rouleau JL, Moye LA, Dagenais GR, Lamas GA, Klein M, Sussex B, Goldman S.** Quantitative Two-dimensional echocardiographic measurements are major predictors of adverse cardiovascular events after acute myocardial infarction. The protective effects of captopril.

Circulation 1994; 89: 68-75.

11. **Taylor G, Humphries J, Mellits E, Pitt B, Schulze R, Griffith L, Achuff S.** Predictors of clinical course, coronary anatomy and left ventricular function after recovery from acute myocardial infarction. *Circulation* 1980; 62(5): 960-70.
12. **White HD, Norris RM, Brown MA, Brandt PW, Whitlock RM, Wild CJ.** Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation* 1987; 76: 44-51.
13. **MacDermott K, Minhas R, Packham C, Skinner J, Squires H, Thomson D, Timmis A, Walsh J, White A, Williams H, Cooper A, Kathoria M.** Secondary prevention in primary and secondary care for patients following a myocardial infarction; This guideline replaces NICE inherited guideline A. NICE, May 2007.
14. **Modern standards and service models, coronary heart disease: chapter three-heart attacks & other acute coronary syndromes.** NHS National Service Frameworks, Mar 2005.
15. **Mandinov L, Eberli FR, Seiler C, Hess OM.** Diastolic heart failure. *Cardiovasc Res* 2000; 45(4): 813-25.
16. **Kirkpatrick JN, Vannan MA, Narula J, Lang RM.** Echocardiography in Heart failure: applications, utility, and new horizons. *J Am Coll Cardiol* 2007; 50: 381-96.
17. **American Heart Association, American Stroke Association.** Heart disease and stroke statistics, 2005.
18. **Krumholz HM, Chen YT, Wang Y, Vaccarino V, Radford MJ, Horwitz RI.** Redictors of readmission among elderly survivors of admission with heart failure. *Am Heart J* 2000; 139: 72-7.
19. **Nieminen MS, Bohm M, Cowie MR, Drexler H, Filippatos GS, Jondeau G, Hasin Y, Lopez-Sendon J, Mebazaa A, Metra M, Rhodes A, Swedberg K.** Guidelines on the diagnosis and treatment of acute heart failure-full text. The task force on acute heart failure of the European Society of Cardiology. European Society of Cardiology 2005.
20. **Sanderson JE, Chan SK, Chan WW, Hung YT, Woo KS.** The aetiology of heart failure in the Chinese population of Hong Kong—a prospective study of 730

consecutive patients. *Int J Cardiol 1995*; 51: 29-35.

21. **Neubauer S.** The failing heart – an engine out of fuel. *N Engl J Med* 2007; 356: 1140-51.
22. **Hunt SA, Abraham WT, Chin MH.** ACC/AHA 2005 guideline update for the diagnosis and management of chronic heart failure in the adult: summary article: a report of the Amercian College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2005; 46: e1-82.
23. **Lang R, Bierig M, Devereaux R, Flachskampf F, Foster E, Pellikka P, Picard M, Roman M, Seward J, Shanewise J.** Recommendations for chamber quantification. J Am Soc Echocardiogr 2005; 18: 1440-63.
24. **Vasan RS, Larson MG, Benjamin EJ, Evans JC, Reiss CK, Levy D.** Congestive heart failure in subjects with normal versus reduced left ventricular ejection fraction: prevalence and mortality in a population-based cohort. *J Am Coll Cardiol* 1999; 33: 1948-55.
25. **Grayburn PA, Appleton CP, DeMaria AN.** Echocardiographic predictors of morbidity and mortality in patients with advanced heart failure. *J Am Coll Cardiol* 2005; 45: 1064-71.
26. **Moss AJ, Zareba W, Hall WJ, Klein H, Wilber DJ, Cannom DS, Daubert JP, Higgins SL, Brown MW, Andrews ML.** Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *N Engl J Med* 2002; 346: 877-83.
27. **Eagle KA, Guyton RA, Davidoff R, Edwards FH, Ewy GA, Gardner TJ, Hart JC, Herrmann HC, Hillis LD, Hutter AM Jr, Lytle BW, Marlow RA, Nugent WC, Orszulak TA, Antman EM, Smith SC Jr, Alpert JS, Anderson JL, Faxon DP, Fuster V, Gibbons RJ, Gregoratos G, Halperin JL, Hiratzka LF, Hunt SA, Jacobs AK, Ornato JP.** ACC/AHA 2004 guideline update for coronary artery bypass graft surgery: summary article. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for Coronary Artery Bypass Graft Surgery). *J Am Coll Cardiol* 2004; 44:1146–310.

28. **Jacobs LD, Salgo IS, Goonewardena S, Weinert L, Coon P, Bardo D, Gerard O, Allain P, Zamorano JL, de Isla LP, Mor-Avi V, Lang RM.** Rapid online quantification of left ventricular volume from real-time three-dimensional echocardiographic data. *Eur Heart J* 2006; 27: 460-8.
29. **Anderson B**. Echocardiography: the normal examination and echocardiographic measurements. *Blackwell Publishing* 2000.
30. **Mor-Avi V, Sugeng L, Lang RM.** Real-time three-dimensional echocardiography: an integral component of the routine echocardiographic examination in adult patients? *Circulation* 2009; 119: 314-29.
31. **Qin JX, Jones M, Shiota T, Greenberg NL, Tsujino H, Firstenberg MS, Gupta PC, Zetts AD, Xu Y, Ping SJ, Cardon LA, Odabashian JA, Flamm SD, White RD, Panza JA, Thomas JD.** Validation of real-time three-dimensional echocardiography for quantifying left ventricular volumes in the presence of a left ventricular aneurysm: in vitro and in vivo studies. *J Am Coll Cardiol* 2000; 36: 900-7.
32. **Arai K, Hozumi T, Matsumura Y, Sugioka K, Takemoto Y, Yamagishi H, Yoshiyama M, Kasanuki H, Yoshikawa J.** Accuracy of measurement of left ventricular volume and ejection fraction by new real-time three dimensional echocardiography in patients with wall motion abnormalities secondary to myocardial infarction. *Am J Cardiol* 2004; 94: 552-8.
33. **Jenkins C, Bricknell K, Hanekom L, Marwick TH.** Reproducibility and accuracy of echocardiographic measurements of left ventricular parameters using real-time three-dimensional echocardiography. *J Am Coll Cardiol* 2004; 44: 878-86.
34. **Nikitin NP, Constantin C, Loh PH, Ghosh J, Lukaschuk EI, Bennett A, Hurren S, Alamgir F, Clark AL, Cleland JG.** New generation 3-dimensional echocardiography for left ventricular volumetric and functional measurements: comparison with cardiac magnetic resonance. *Eur J Echocardiogr* 2006; 7: 365-72.
35. **Tighe DA, Rosetti M, Vinch CS, Chandok D, Muldoon D, Wiggin B, Dahlberg ST, Aurigemma GP.** Influence of image quality on the accuracy of real time 3 dimensional echocardiography to measure left ventricular volumes in unselected

patients: a comparison with gated-SPECT imaging. *Echocardiography* 2007; 24: 1073-80.

36. **Pouleur AC, le Polain de Waroux JB, Pasquet A, Gerber BL, Gerard O, Allain P, Vanoverschelde JL.** Assessment of left ventricular mass and volumes by three-dimensional echocardiography in patients with or without wall motion abnormalities: comparison against cine magnetic resonance imaging. *Heart* 2008; 94: 1050-7.
37. **Chan J, Jenkins C, Khafagi F, Du L, Marwick TH.** What is the optimal clinical technique for measurement of left ventricular volume after myocardial infarction? A comparative study of 3-dimensional echocardiography, single photon emission computed tomography, and cardiac magnetic resonance imaging. *J Am Soc Echocardiogr 2006*; 19: 192–201.
38. **Chuang ML, Hibberd MG, Salton CJ, Beaudin RA, Riley MF, Robert AP, Douglas PS, Manning WJ**. Importance of imaging method over imaging modality in noninvasive determination of left ventricular volumes and ejection fraction: Assessment by two- and three-dimensional echocardiography and magnetic resonance imaging. *J Am Coll Cardiol* 2000; 35: 477-84.
39. **Gopal AS, Keller AM, Rigling R, King DL Jr, King DL**. Left ventricular volume and endocardial surface area by three-dimensional echocardiography: comparison with two-dimensional echocardiography and nuclear magnetic resonance imaging in normal subjects. *J Am Coll Cardiol* 1993; 22: 258-70.
40. **Nosir YFM, Fioretti PM, Vletter WB, Boersma E, Salustri A, Postma JT, Reijs AEM, Ten Cate FJ, Roelandt JRTC**. Accurate measurement of left ventricular ejection fraction by three-dimensional echocardiography. A comparison with radionuclide angiography. *Circulation* 1996; 94: 460-6.
41. **Nosir YFM, Lequin MH, Kasprzak JD, Domburg RTV, Vletter WB, Yao J, Stoker J, Ten Cate FJ, Roelandt JRTC.** Measurements and day-to-day variabilities of left ventricular volumes and ejection fraction by three-dimenstional echocardiography and comparison with magnetic resonance imaging. *Am J Cardiol* 1998; 82: 209-14.

42. **Nosir YFM, Fioretti PM, Vletter WB, Boersma E, Salustri A, Postma JT, Reijs AE, Ten Cate FJ, Roelandt JR.** Accurate measurement of left ventricular ejection fraction by three-dimensional echocardiography. A comparison with radionuclide angiography. *Circulation* 1996; 94: 460-6.
43. **Kawai J, Tanabe K, Wang CL, Tani T, Yagi T, Shiotani H. Morioka S.** Comparison of left atrial size by freehand scanning three-dimensional echocardiography and two-dimensional echocardiography. *Eur J Echo* 2004; 5: 18-24.
44. **Nosir YFM, Stoker J, Kasprzak JD, Lequin MH, Tencate FJ, Roelandt JRTC.** Paraplane analysis from precordial three-dimensional echocardiographic data sets for rapid and accurate quantification of left ventricular volume and function : A comparison with magnetic resonance imaging. *Am Heart J* 1999; 137: 134-43.
45. **Gutiérrez-Chico JL, Zamorano JL, Isla LPD, Orejas M, Almeria C, Rodrigo JL, Ferreiros J, Serra V, Macaya C.** Comparison of left ventricular volumes and ejection fractions measured by three-dimensional echocardiography versus by two-dimensional echocardiography and cardiac magnetic resonance in patients with various cardiomyopathies. *Am J Cardiol* 2005; 95: 809-13.
46. **Hibberd MG, Chuang ML, Beaudin RA, Riley MF, Mooney MG, Fearnide JT, Manning WJ, Douglas PS.** Accuracy of three-dimensional echocardiography with unrestricted selection of imaging planes for measurement of left ventricular volumes and ejection fraction. *Am Heart J* 2000; 140: 469-75.
47. **Chuang ML, Beaudin RA, Riley MF, Mooney MG., Manning WJ, Douglas PS, Hibberd MG.** Three-dimensional echocardiographic measurement of left ventricular mass: Comparison with magnetic resonance imaging and two-dimensional echocardiographic determinations in man. *Int J Card Imag* 2000; 16: 347-57.
48. **Remin L, Chen SYJ, Mor-Avi V, Bednarz J, Chin-Tu C, Lang RM.** Three-dimensional reconstruction of LV endocardial surfaces from echocardiographic images using deformable shell models. *Comp in Cardiology* 1996; 8: 697-700.

49. **Gopal AS, Shen Z, Sapin PM, Keller AM, Schnellbaecher MJ, Leibowitz DW, Akinboboye OO, Rodney RA, Blood DK, King DL.** Assessment of cardiac function by three-dimensional echocardiography compared with conventional non-invasive methods. *Circulation* 1995;92:842-53.
50. **Kühl HP, Schreckenberg M, Rulands D, Katoh M, Schäfer W, Schummers G, Bücker A, Hanrath P, Franke A.** High-resolution transthoracic real-time three-dimensional echocardiography: quantitation of cardiac volumes and function using semi-automatic border detection and comparison with cardiac magnetic resonance imaging. *J Am Coll Cardiol* 2004; 43: 2083-90.
51. **Wahba G.** Spline models for observational data. In: CBMS-NSF Regional Conference Series in Applied Mathematics 59. *Philadelphia, PA: Society of Industrial and Applied Mathematics (SIAM)* 1990.
52. **Kühl HP, Franke A, Merx MW, Hoffmann R, Puschmann D, Hanrath P.** Rapid quantification of left ventricular function and mass using transesophageal three-dimensional echocardiography: validation of a method that uses long-axis cutplanes. *Eur J Echocardiography* 2000; 1: 213–21.
53. **Kühl HP, Hanrath P, Franke A.** M-Mode Echocardiography overestimates left ventricular mass in patients with normal left ventricular shape: a comparative study using three-dimensional echocardiography. *Eur J Echocardiography* 2003; 4: 313-9.
54. **Osher S, Sethian JA.** Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J Comput Phys* 1988; 79: 12–49.
55. **Malladi R, Sethian JA, Vemuri BC.** Shape modelling with front propagation: a level set approach. *IEEE Trans Pattern Anal Mach Intell* 1995; 17: 158–75.
56. **Caiani EG, Corsi C, Sugeng L, MacEneaney P, Weinert L, Mor-Avi V, Lang RM.** Improved quantification of left ventricular mass based on endocardial and epicardial surface detection with real time three dimensional echocardiography. *Heart* 2006; 92: 213-9.
57. **Gerard O, Billon AC, Rouet JM, Jacob M, Fradkin M, Allouche C.** Efficient model-based quantification of left ventricular function in 3-D echocardiography.

IEEE Trans Med Imaging 2002; 21: 1059-68.

58. **Corsi C, Lang RM, Veronesi F, Weinert L, Caiani EG, MacEneaney P, Lamberti C, Mor-Avi V.** Volumetric quantification of global and regional left ventricular function from real-time three-dimensional echocardiographic images. *Circulation* 2005; 112: 1161-70.
59. **Chukwu EO, Katz AS, Toole RS, Schapiro W, Reichek N, Gopal AS.** Prerequisites for avoiding manual boundary editing using semi-automated left ventricular quantitation by realtime three-dimensional echocardiography. *J Am Coll Cardiol* 2007; 49: 148A.
60. **Soliman Oil, Krenning BJ, Geleijnse ML, Nemes A, van Geuns R, Baks T, Anwar AM, Galema TW, Vletter WB, Gate FJT.** A comparison between QLAB and TomTec full volume reconstruction for real time three-dimensional echocardiographic quantification of left ventricular volumes. *Echocardiography* 2007; 24(9): 967-74.
61. **Hansegard J, Urheim S, Lunde K, Malm S, Rabben IS.** Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography. *Cardiovasc Ultrasound* 2009; 7: 18.
62. **Sanmarco ME, Bartle SH.** Measurement of left ventricular volume in the canine heart by biplane angiocardiography: accuracy of the method using different model analogies. *Circ Res* 1966; 19: 11-7.
63. **Pentland A, Sclaroff S.** Closed-form solutions for physically based shape modeling and recognition. *IEEE Trans Patt Anal and Mach Intell* 1991; 13: 715-29.
64. **William KP.** Digital image processing, second edition. *A Wiley-Interscience Publication, John Wiley & Sons, INC* 1991.
65. **Nielsen PMF, Grice IJ, Smaill BH, Hunter PJ.** Mathematical model of geometry and fibrous structure of the heart. Am J Physiol 1991; 260: 1365-78.
66. **Zienkiewicz OC, Morgan K.** Finite elements and approximation. *Dover Publications* 2006
67. **Bland J, Altman D.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; i: 307-10.

68. **Madsen EL, Frank GR, Dong F.** Liquid or solid ultrasonically tissue-mimicking materials with very low scatter. *Ultrasound in Med & Biol* 1998; 24: 535-42.
69. **Al-Nassar YN, Al-Jalal AM, Khan MA, Al-Kaabi.** Functional dependence of ultrasonic speed in water on salinity and temperature. *NDT.net* 2006; 11(6).
70. **Pattynama PM, Lamb HJ, van der Velde EA, van der Wall EE, de Roos A.** Left ventricular measurements with cine and spin-echo MR imaging: a study of reproducibility with variance component analysis. *Radiology* 1994; 187: 261-8.
71. **Bellenger NG, Davies LC, Francis JM, Coats AJ, Pennell DJ.** Reduction in sample size for studies of remodeling in heart failure by the use of cardiovascular magnetic resonance. *J Cardiovasc Magn Reson* 2000; 2: 271-8.
72. **van Geuns RJ, Baks T, Gronenschild EH, Aben JP, Wielopolski PA, Cademartiri F, de Feyter PJ.** Automatic quantitative left ventricular analysis of cine MR images by using three-dimensional information for contour detection. *Radiology* 2006; 240: 215-21.
73. **Codella NC, Weinsaft JW, Cham MD, Janik M, Prince MR, Wang Y.** Left ventricle: automated segmentation by using myocardial effusion threshold reduction and intravoxel computation at MR imaging. *Radiology* 2008; 248: 1004-12.
74. **Hudsmith LE, Petersen SE, Francis JM, Robson MD, Neubauer S.** Normal human left and right ventricular and left atrial dimensions using steady state free precession magnetic resonance imaging. *J Cardiovasc Magn Reson* 2005; 7: 775-82.
75. **Koskenvuo JW, Karra H, Lehtinen J, Niemi P, Pärkkä J, Knuuti J, Hartiala JJ.** Cardiac MRI: accuracy of simultaneous measurement of left and right ventricular parameters using three different sequences. Clin Physiol Funct Imaging 2007; 27(6): 385-93.
76. **Hoffmann R, Bardeleben SV, Cate FT, Borges AC, Kasprzak J, Firschke C, Lafitte S, Al-Saadi N, Kuntz-Hehner S, EngelHardt M, Becher H, Vanoverschelde JL.** Assessment of systolic left ventricular function: a multi-centre comparison of cineventriculography, cardiac magnetic resonance imaging, unenhanced and contrast-enhanced echocardiography. *Eur Heart J* 2005; 26:

607-16.

77. **Argyle E.** Techniques for edge detection. *Proc. IEEE* 1971; 59 (2): 285-7.
78. **Macleod IDG.** Picture processing and psychopictorics. *New York Academic Press* 1970.
79. **Macleod IDG.** Comments on techniques for edge detection. *Proc. IEEE* 1972; 60: 344.