Université de Montréal

Semiautomated 3D Liver Segmentation Using Computed Tomography and Magnetic Resonance Imaging

par
Akshat Gotra, MD

Département de radiologie
Faculté de médecine

Mémoire présenté à la Faculté de médecine
en vue de l'obtention du grade de maîtrise
en sciences biomédicales
option recherche clinique biomédicale

Août 2015

Résumé

Le foie est un organe vital ayant une capacité de régénération exceptionnelle et un rôle crucial dans le fonctionnement de l'organisme. L'évaluation du volume du foie est un outil important pouvant être utilisé comme marqueur biologique de sévérité de maladies hépatiques. La volumétrie du foie est indiquée avant les hépatectomies majeures, l'embolisation de la veine porte et la transplantation.

La méthode la plus répandue sur la base d'examens de tomodensitométrie (TDM) et d'imagerie par résonance magnétique (IRM) consiste à délimiter le contour du foie sur plusieurs coupes consécutives, un processus appelé la «segmentation».

Nous présentons la conception et la stratégie de validation pour une méthode de segmentation semi-automatisée développée à notre institution. Notre méthode représente une approche basée sur un modèle utilisant l'interpolation variationnelle de forme ainsi que l'optimisation de maillages de Laplace. La méthode a été conçue afin d'être compatible avec la TDM ainsi que l' IRM.

Nous avons évalué la répétabilité, la fiabilité ainsi que l'efficacité de notre méthode semi-automatisée de segmentation avec deux études transversales conçues rétrospectivement. Les résultats de nos études de validation suggèrent que la méthode de segmentation confère une fiabilité et répétabilité comparables à la segmentation manuelle. De plus, cette méthode diminue de façon significative le temps d'interaction, la rendant ainsi adaptée à la pratique clinique courante.

D'autres études pourraient incorporer la volumétrie afin de déterminer des marqueurs biologiques de maladie hépatique basés sur le volume tels que la présence de stéatose, de fer, ou encore la mesure de fibrose par unité de volume.

Mots-clés : Foie, segmentation hépatique, volumétrie, imagerie 3D, validation, planification préopératoire, tomodensitométrie, imagerie par résonance magnétique, humains.

Abstract

The liver is a vital abdominal organ known for its remarkable regenerative capacity and fundamental role in organism viability. Assessment of liver volume is an important tool which physicians use as a biomarker of disease severity. Liver volumetry is clinically indicated prior to major hepatectomy, portal vein embolization and transplantation.

The most popular method to determine liver volume from computed tomography (CT) and magnetic resonance imaging (MRI) examinations involves contouring the liver on consecutive imaging slices, a process called "segmentation". Segmentation can be performed either manually or in an automated fashion.

We present the design concept and validation strategy for an innovative semiautomated liver segmentation method developed at our institution. Our method represents a model-based approach using variational shape interpolation and Laplacian mesh optimization techniques. It is independent of training data, requires limited user interactions and is robust to a variety of pathological cases. Further, it was designed for compatibility with both CT and MRI examinations.

We evaluated the repeatability, agreement and efficiency of our semiautomated method in two retrospective cross-sectional studies. The results of our validation studies suggest that semiautomated liver segmentation can provide strong agreement and repeatability when compared to manual segmentation. Further, segmentation automation significantly shortens interaction time, thus making it suitable for daily clinical practice.

Future studies may incorporate liver volumetry to determine volume-averaged biomarkers of liver disease, such as such as fat, iron or fibrosis measurements per unit volume. Segmental volumetry could also be assessed based on subsegmentation of vascular anatomy.

Keywords : Liver, Segmentation, Volumetry, 3D Imaging, Validation, Preoperative planning, Computed Tomography, Magnetic Resonance Imaging, Humans.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations and Symbols

3D	=	*Three-dimensional*
ASD	=	*Average symmetric surface distance*
BSA	=	*Body surface area*
CI	=	*Confidence interval*
CT	=	*Computed tomography*
DICOM	=	*Digital Imaging and Communications in Medicine*
FLASH	=	*Fast low-angle shot sequence*
FLR	=	*Future liver remnant*
GRE	=	*Gradient-recalled echo*
HCC	=	*Hepatocellular carcinoma*
ICC	=	*Intraclass correlation coefficient*
IRM	=	*Imagerie par résonance magnétique*
IVC	=	*Inferior vena cava*
LAVA	=	*Liver acquisition with volume acceleration*
LIO	=	*Laboratoire de recherche en imagerie et orthopédie*
LI-RADS	=	*Liver Imaging Reporting and Data System*
MDCT	=	*Multi-detector computed tomography*
MPR	=	*Multi-planar reconstruction*
MRI	=	*Magnetic resonance imaging*
MRS	=	*Magnetic resonance spectroscopy*
MSD	=	*Maximum symmetric surface distance*
NAFLD	=	*Nonalcoholic fatty liver disease*
NASH	=	*Nonalcoholic steatohepatitis*
PDFF	=	*Proton density fat fraction*
PVE	=	*Portal vein embolization*
RBF	=	*Radial basis function*
RMSD	=	*Root mean square symmetric surface distance*
ROI	=	*Region of interest*

RVD	=	*Relative volume difference*
SD	=	*Standard deviation*
SSM	=	*Statistical shape models*
TDM	=	*Tomodensitométrie*
TLFI	=	*Total liver fat index*
TLV	=	*Total liver volume*
VOE	=	*Volumetric overlap error*
VIBE	=	*Volumetric interpolated breath-hold examination*
VOI	=	*Volume of interest*

For my mother and father who have selflessly sacrificed to allow us to achieve.

Acknowledgments

I am immensely grateful to my research supervisor and mentor, Dr. An Tang. At a time of uncertainty you provided solid ground in the form of a carefully considered and meticulously organized research program. I thank you for the hours spent vigilantly reviewing my manuscripts and presentations despite a burgeoning research career and numerous other commitments. You were an exceptional teacher during these formative years.

I feel tremendously privileged to have worked with Gabriel Chartrand who was largely responsible for developing the liver segmentation software described in this dissertation. Many of the renderings seen on these pages are a result of his visual artistry. Without your help, these studies would not have been possible. I am glad I was able to work with you during the early stages of what will surely be a brilliant career.

I am thankful to have met and worked with members of the Laboratoire de recherche en imagerie et orthopédie (LIO) in Montreal, Quebec. In particular, thanks to Dr. Jacques A. de Guise and Ramnada Chav for providing resources and guidance while making me feel welcome in the laboratory.

I am grateful to Dr. Jeffrey Chankowsky and Dr. Benoît Gallix from McGill University for being supportive of my Masters degree requirements during formal residency training. Your continued clinical guidance and mentorship is greatly appreciated.

I am indebted to Dr. Gilles Soulez, Dr. Samuel Kadoury and Dr. Gallix for serving as jury members for my Masters thesis.

Finally I would like to thank my parents, Amar Gotra and Kusum Nehru, and my brother Anuj for their continued love and support. This dissertation is dedicated to you.

1 Introduction

This dissertation introduces an original software approach to 3-dimensional liver segmentation from multi-planar imaging. The sophistication of modern imaging techniques allow the physician to visualize human anatomy in an unparalleled fashion. Images generated from modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) are of higher quality, acquired faster and provide wider coverage than in prior years. This has led to a surge in the amount of data that must be scrutinized during each imaging examination.

The scientific community has countered with computer-aided tools, which assist the physician in efficiently extracting relevant information from each imaging exam. One such image processing tool consists of delineation of a region of interest on CT and MRI images for volumetric analysis. This process is called *segmentation* and its optimization has been a research focus over decades within the biomedical field.

Organ segmentation has a variety of medical applications ranging from surgical simulation to radiotherapy planning. The medical literature describes numerous segmentation methods and algorithms adapted for different target organs. The liver remains one of the most challenging organs to efficiently segment due to, among other reasons, its highly variable shape and close proximity to other organs and tissues.

Segmentation has traditionally been performed by manually outlining the liver contour on each individual CT or MRI image. Given the hundreds of images acquired on each examination, the manual method is too time-consuming and cumbersome for daily clinical use. Segmentation automation is essential to improve time-efficiency without sacrificing volumetric accuracy and repeatability.

The primary aim of this thesis is to describe the development and validation of an innovative semiautomated liver segmentation method developed at our

institution. We wish to assess the method's accuracy, repeatability and efficiency in segmenting both CT and MRI examinations.

1.1 The Liver

"He bound devious Prometheus with inescapable harsh bonds, fastened through the middle of a column, and he inflicted on him a long-winged eagle, which ate his immortal liver, but it grew as much in all at night as the long-winged bird would eat all day."

- Hesiod's Theogony (1)

Prometheus is a Titan from Greek mythology known for gifting mankind with fire stolen from Mount Olympus. As punishment, he was chained and condemned to having an eagle eternally feast on his liver. The legend specifies that his "immortal liver" would grow each night, suggesting that the ancient Greeks were aware of the liver's regenerative capacity (2). The metaphorical significance of the liver has been highlighted in other texts where it is equated with the eternal soul and with intelligence (3).

Evolutionary events have indeed imparted the liver with a remarkable capacity to regenerate following loss of mass in vertebrates (4). This is likely a result of the essential bodily functions the liver performs and its fundamental role in organism viability. This regenerative capacity had been mythologized since ancient times.

While liver divination may not be as popular today, the liver's role as a vital organ is definitely recognized. The liver is the largest human organ and is located in the right upper quadrant of the abdomen, below the diaphragm, adjacent to the stomach and overlying the gallbladder. The hepatocyte is the basic metabolic cell of the liver. Millions of hepatocytes constitute the lobule, which is the basic functional unit. The liver is known to have a role in roughly 500 vital bodily functions (5). Though out of the scope of this dissertation, some of these

roles include: digestion (bile production), metabolism (i.e. proteins, carbohydrates, lipids), synthesis (i.e. albumin, hormones), storage (i.e. glycogen, vitamins), detoxification (i.e. alcohol, drugs) and immunity (5).

The liver is closely associated with three major blood vessels: the hepatic artery, the portal vein and the inferior vena cava (IVC). The hepatic artery supplies the liver with oxygenated blood stemming from the abdominal aorta. The portal vein supplies the liver with nutrient-rich blood derived from gastro-intestinal organs (i.e. stomach, small bowel and colon). These two vessels along with the common bile duct enter the liver through a deep fissure at its inferior surface known as the *porta hepatis*, or simply, the liver hilum.

Every liver lobule is supplied by a tributary of the hepatic artery, portal vein and common bile duct. Each lobule is drained by a branch of the hepatic veins, and subsequently, the IVC. The IVC is a retroperitoneal structure that runs along the right side of the vertebral column, lateral to the aorta **(Figure 1.1)**. The right, middle and left hepatic veins drain into the IVC at the level of thoracic vertebrae eight.

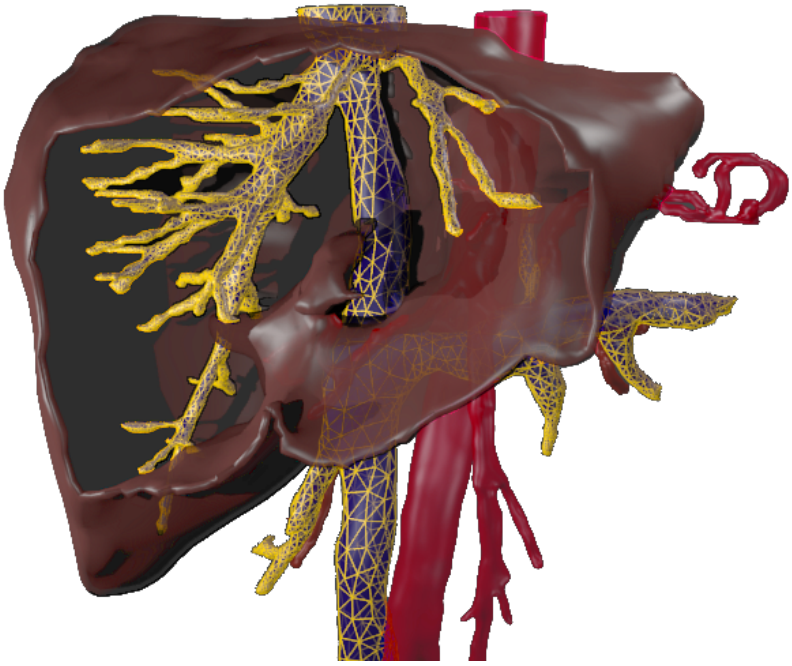

Figure 1.1: 3D rendering of the liver and associated vascular structures
Basic anatomical structure of the liver and spatial relationship with major vascular structures; the **aorta** and branches (red) and the **IVC** and hepatic veins (blue/yellow). The three hepatic veins are pictured draining into the IVC, a retroperitoneal structure. This usually occurs at the level of thoracic vertebrae eight. Image courtesy of Gabriel Chartrand.

Couinaud classification

Claude Couinaud introduced the Couinaud classification system in 1957 (6). The system describes functional liver anatomy by dividing the liver into eight independent segments. Each segment has its own respective vascular inflow, outflow, biliary and lymphatic drainage. This segmental classification is of particular importance during surgical planning as independent segments may be resected without affecting the remaining ones. Further, using a common classification system allows for simple communication between physicians from different specialties when describing focal liver lesions (7). The eight functional segments are separated based on vascular anatomy as follows:

1. The horizontal plane of the portal vein bifurcation divides the liver into upper and lower sections.
2. The vertical plane of the middle hepatic vein divides the liver into two halves, establishing the right liver and the left liver.
3. The right hepatic vein divides the right liver into anterior and posterior segments.
4. The left hepatic vein divides the left liver into medial and lateral segments.

Though Couinaud initially described segments based on portal vein branching, the working system is actually based on the three hepatic veins (7). Segment I (caudate lobe) is formed by the liver tissue located between the portal bifurcation and the IVC. The remainder of the segments (II to VIII) are numbered in a clockwise fashion **(Figure 1.2)**.

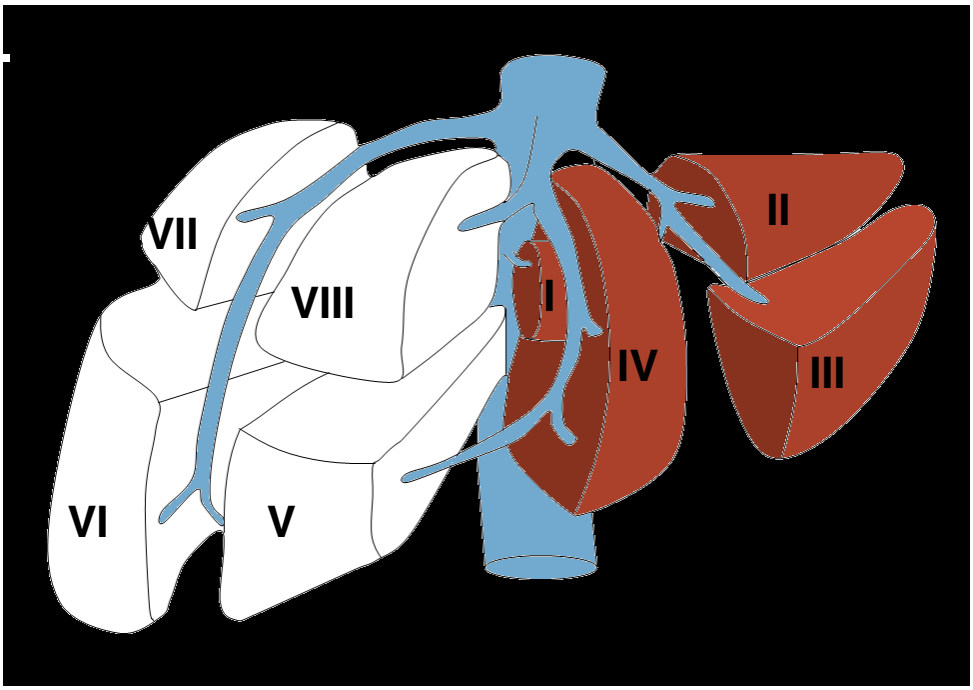

Figure 1.2: Couinaud classification of liver segments.
The portal vein bifurcation establishes a horizontal plane dividing the liver into upper and lower sections. The middle hepatic vein divides the liver into the right (white) and left liver (orange). Segments II-VIII are numbered clockwise beginning at segment II (8).

1.2 Liver Volumetry

The liver is subject to variety of diverse pathology which is often debilitating to the individual patient. Medical and surgical treatments offered for these pathologies range from pharmaceutical management to surgical excision or transplant. Assessment of liver volume represents a basic tool which physicians often use for diagnosis of diffuse or metastatic liver disease that may present as hepatomegaly.

The need for accurate liver volumetry has been expressed in both medical and surgical contexts. Medically, the liver volume is known to be an important prognostic indicator in compensated cirrhosis (9) and fulminant liver failure (10).

Linguraru et al. also recently described the clinical significance of volumetry in assessing hepatomegaly (11).

In the surgical context, liver volumetry is essential during hepatectomy planning to ensure residual liver regeneration and prevent post-operative hepatic failure (12). The importance of volumetry in surgical planning for orthotopic liver transplantation (13) and living-donor liver transplantation (14-17) has been emphasized. Other local liver interventions including radiotherapy, radio-frequency ablation and cryo-ablation may also require liver volume assessment prior to treatment (18), especially if the patient has previously undergone hepatectomy.

The role of liver volume as a potential biomarker for liver disease and the established clinical indications for volumetry are described in **Sections 2.2** and **2.3**, respectively.

Traditionally, liver size has been estimated crudely. Medical students are taught early in training how to inspect, palpate, percuss and auscultate the abdomen in order to elicit pathological involvement. The standard physical exam includes a measurement of liver size using manual palpation and percussion to identify the liver margins (19).

Liver size varies greatly and is dependent on a variety of factors including: age, body size, shape, underlying pathology and more importantly, the examination technique used (i.e. palpation, percussion or radiographic). In 1977, the mean liver size was thought to be 7 cm for women and 10.5 cm for men, a liver span of two to three centimeters smaller or larger by physical exam being considered abnormal (20). Techniques measuring one-dimensional length are often criticized as they are rudimentary and subject to significant variability depending on underlying liver shape, pathology and degree of lung inflation. Further, livers of different shape and volume may have the same cranial-caudal length **(Figure 1.3)**.

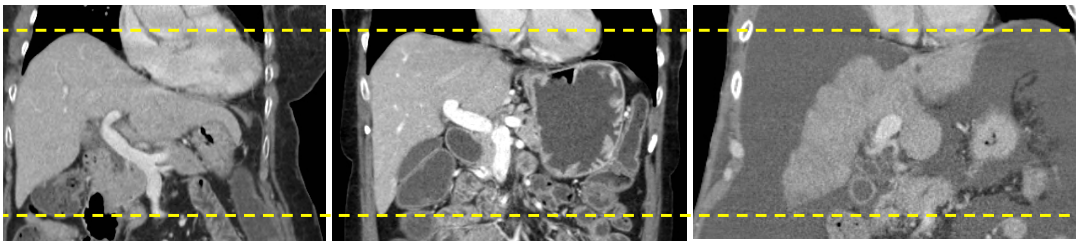

Figure 1.3: Variability in liver shape and size
Livers of different shape and volume may have the same cranial-caudal length, as demonstrated with these three examples. This observation highlights the limitation of reporting a one-dimensional measures of length, a well-entrenched practice, as a surrogate measure of liver volume. Image courtesy of Dr. An Tang.

The advent of modern cross-sectional imaging techniques provided additional tools to estimate liver volume. Assessment of whole liver volume using CT was initially demonstrated in 1979 as a means to compare the effect of portosystemic shunts on hepatic structure and function (21). Heymsfield et al. measured the volume of a cadaveric liver using CT images and showed a discrepancy of less than 5% with volume obtained from the water displacement method (21).

Today, CT and MRI are commonly used to meet the clinical need of accurate estimation of liver volume. Use of CT is often preferred due to its easier accessibility, high spatial resolution, robustness and short acquisition time (14, 22, 23). MRI offers multiple contrast mechanisms and ability to simultaneously assess vascular anatomy, biliary anatomy and liver parenchymal pathology (24). Further, MR imaging minimizes the risk of radiation exposure and nephrotoxicity which are concerns for CT imaging (24, 25).

The most popular method to determine liver volume from CT and MRI images involves contouring the liver outline on consecutive imaging slices, a process called “segmentation”. Segmentation refers to delineation of a region of interest from the background.

1.3 Liver Segmentation

In vivo assessment of liver volume is problematic as it is impossible to directly obtain an exact measurement. As explanting an organ from a living being would be unrealistic, indirect measurements via imaging post-processing tools are sought. Segmentation software indirectly measures volume by identifying the number of voxels, or the smallest distinguishable box-shaped parts of a 3D space, belonging to an organ of interest. The volumetric error is thus directly proportional to the error associated with identification of voxels.

A common approach used in image segmentation consists of demarcating the contours of a structure of interest to identify the number of enclosed voxels. Despite significant technical advances in the field of image processing, this segmentation is often performed manually in the clinical setting **(Figure 1.4)**. The manual segmentation of a liver from CT or MRI images must be performed on each axial slice and can take an image analyst anywhere from 30 to 90 minutes (26). It is time-consuming, cumbersome and expensive and thus not ideal for busy clinical practice.

Figure 1.4: Manual segmentation.
Three selected axial images from a contrast enhanced CT examination demonstrating segmentation of liver contours. Segmentation is performed to enclose the voxels belonging to a structure of interest and indirectly measure volume.

Alternatively, liver segmentation methods requiring minimal to no user input have been a research focus in the field of biomedical engineering for decades. Though numerous studies have proposed semi- or fully-automated segmentation methods, these have not necessarily translated to routine clinical use (27). Limited clinical validation studies for these methods, rather than lack of technical ingenuity, are cited as the cause of this slow adaptation by the medical community (28).

In their study, Udupa et al. classified the weaknesses of image segmentation algorithm evaluation frameworks into two categories: related to available resources or related to the employed methodology (28). Reasons thought to limit the performance of segmentation algorithms include: small sample sizes, data sets not reflective of clinical problems, inappropriate ground truth for comparison and poorly defined performance metrics (27, 28).

In order to overcome these methodological weaknesses, a validation framework for a novel automated segmentation method should include, at minimum, the following elements (28):

1. Use of a valid reference standard.
2. Datasets for validation which are reflective of actual clinical practice; real cases rather than ideal cases.
3. Clear metrics for measurement of segmentation precision, accuracy, efficiency and error.
4. Comparison of metrics for each method using effective statistical tools.

We attempted to incorporate these defined elements in the validation of our novel segmentation method.

1.4 Thesis Structure

This dissertation has been written in such a manner that it is relevant to a broad range of readers including medical students, radiologists, hepatologists, hepatobiliary surgeons and biomedical engineers.

Figure 1.5 provides a roadmap to the material covered in various sections. Medical students should read this document sequentially. Radiologists will be most interested in the figures and tables accompanying the text and may peruse these in sequential order to find areas of interest. Physicians with intimate knowledge about liver disease and clinical indications for liver volumetry may proceed directly to Section 3. Those specifically interested in the clinical validation steps should focus on **Sections 4** and **5**.

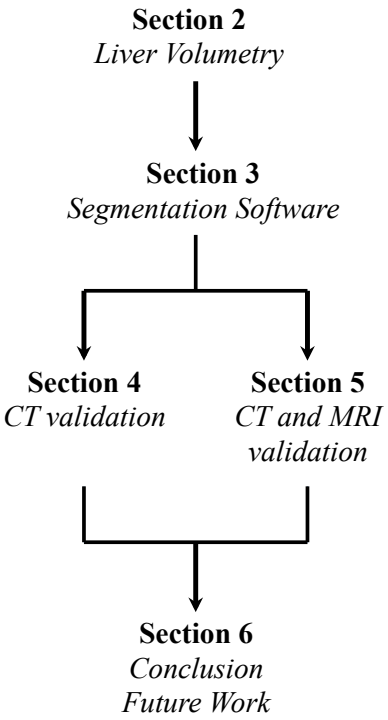

Figure 1.5: Dissertation roadmap.

Section Descriptions

• **Section 2** summarizes the most common liver diseases currently affecting those living in the Western world. Liver volume is defined as a biomarker which can be used in the management and treatment of these diseases. The most common clinical indications for performing liver volumetry are reviewed. Finally, manual and automated liver volumetry methods are introduced.

• **Section 3** introduces the design concept and workflow of the semiautomated liver segmentation software developed at our institution. The three main steps required to perform segmentation: initialization, optimization and correction are described in detail. The final section describes a proof of concept study where the multi-modality versatility of the segmentation method was tested using CT and MRI datasets.

• **Section 4** describes a retrospective, cross-sectional study which evaluated our segmentation method on patients who underwent contrast-enhanced CT prior to major hepatectomy between October 2006 and April 2009. This study was an initial validation step of our method for CT-based imaging.

• **Section 5** describes a retrospective, cross-sectional study which evaluated our segmentation method on subjects who required preoperative evaluation with both CT and MRI within two weeks between January 2010 and March 2013. This study compared the results obtained from semiautomated segmentation of CT and MRI images.

• **Section 6** summarizes the lessons learned during this research process, the challenges ahead and points to future research directions.

2 Liver Volumetry

2.1 Liver Diseases

The liver is intimately involved in most vital processes taking place within the human body. Consequently, a variety of chronic infections and diseases can pathologically affect it. In March 2013 the Canadian Liver Foundation commissioned a report entitled "Liver Disease in Canada, a Crisis in the Making", which assessed the extent of liver disease affecting the Canadian population (29). The report estimates that one in ten Canadians have some form of liver disease and the related death rate has risen nearly 30% over eight years (29).

More than 95% of all deaths from liver disease are attributable to either viral hepatitis, alcoholic liver disease, non-alcoholic fatty liver disease (NAFLD), cirrhosis or hepatocellular carcinoma (HCC). These disease processes ultimately result in the need for medical treatment, liver transplantation or hepatectomy and represent important global health concerns.

2.1.1 Viral Hepatitis

Viral hepatitis refers to liver inflammation and injury secondary to viral infection. Of the offending viruses, hepatitis B and C most commonly infect the liver causing extensive mortality and morbidity (29). Both are blood-borne infections which spread through close contact with infected body fluids.

Hepatitis B is a double-stranded DNA virus of the *Hepadnaviridae* family. By definition, an acute hepatitis B infection lasts six months or less without any permanent damage to hepatocytes and with development of future immunity. A chronic hepatitis B infection lasts longer than 6 months and is usually life-long. Untreated chronic infection may eventually lead to cirrhosis in 15-20% of cases (29). Primary liver cancer develops with an incidence of 0.2-0.6% in the non-cirrhotic/hepatitis B population and 5-8% in the cirrhotic/hepatitis B population

(29). Chronic hepatitis B is highly prevalent (5-12%) amongst Canadian immigrant populations (30, 31). Treatment of chronic infection with anti-viral medications can improve liver function, reduce progression of fibrosis and cirrhosis and reduce the risk of HCC development. All Canadian provinces have instituted either neonatal or adolescent universal vaccination programs against hepatitis B (29).

Hepatitis C is single-stranded RNA virus of the *Flaviviridae* family. Acute infection with hepatitis C is rarely symptomatic. Failure to spontaneously clear the infection leads to chronic infection, with possible progression to cirrhosis and liver failure. Liver failure secondary to chronic hepatitis C infection represents the most common indication for liver transplantation in Canada (29). At time of reporting, hepatitis C has a peak prevalence in middle-aged individuals aged 30-59 years. Hepatitis C is considered a curable disease with anti-viral treatment regimens including interferon alpha, ribonucleic acid analogs (Ribavirin) and protease inhibitors. Cure rates range from 60-75% depending on the genotype and treatment length (29). No vaccination strategies currently exist against Hepatitis C, although several are in development (32).

2.1.2 Alcoholic Liver Disease

Alcoholic liver disease results from excessive alcohol consumption and represents an important cause of worldwide morbidity and mortality. Alcohol causes damage to hepatocytes directly as a toxic substance and indirectly by promoting hepatitis C infection and insulin resistance with subsequent fatty liver disease (29).

Alcoholic liver disease exists in two main forms: acute alcoholic hepatitis and alcoholic cirrhosis. Alcoholic hepatitis is usually characterized by acute clinical and biochemical evidence of liver failure. In the context of pre-existing cirrhosis, this condition can be fatal. Lesser but prolonged drinking can lead directly to

alcoholic cirrhosis, which has a very poor prognosis as 47% of those afflicted die within 5 years (28).

2.1.3 Non-Alcoholic Fatty Liver Disease

Non-Alcoholic Fatty Liver Disease (NAFLD) is an infiltrative disease of the liver associated with obesity and type 2 diabetes. NAFLD is linked to a group of conditions collectively termed the "metabolic syndrome". This syndrome is characterized by a resistance to insulin that favors the intracellular accumulation of fatty acids and triglycerides (33). Fatty acids are known to cause oxidative stress by stimulating stellate cells responsible for hepatic injury and fibrosis, eventually leading to cirrhosis (34). NAFLD is currently the most common liver disease affecting the Canadian population (29).

This disease evolves through a spectrum of three main stages. The initial stage is *steatosis*, a broad term denoting fat accumulation within the liver without significant inflammation or fibrosis. The second stage is non-alcoholic steatohepatitis (NASH), where fat accumulation is associated with inflammatory changes and scarring. Of the patients who evolve to NASH, half will develop liver fibrosis, whereas nearly 20% will experience either cirrhosis or liver failure (35, 36). Cirrhosis represents the third stage of the disease. The stages and complications of NAFLD are outlined in **Figure 2.1**.

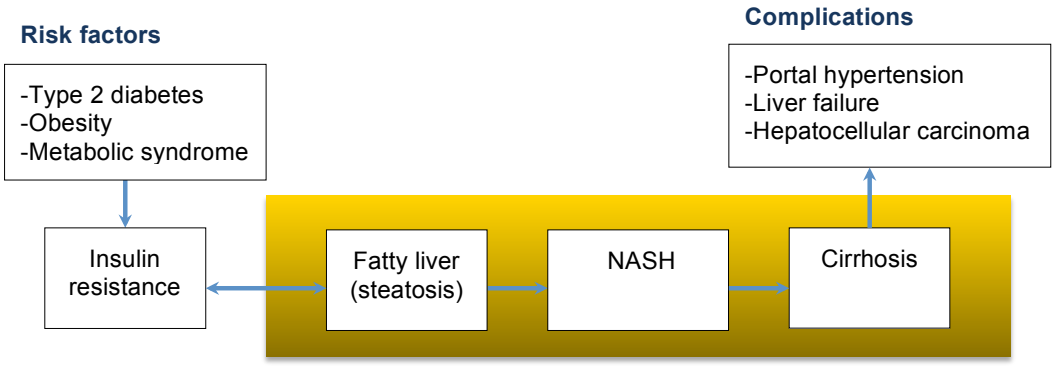

Figure 2.1: Pathophysiology, risk factors and potential complications of NAFLD The main stages of NAFLD (i.e. steatosis, NASH and cirrhosis) are outlined in yellow. The major risk factors for NAFLD (i.e. type 2 diabetes, obesity and metabolic syndrome) are associated with insulin resistance. Insulin resistance promotes the intracellular accumulation of fatty acids which can eventually lead to hepatic injury, fibrosis and cirrhosis. Cirrhosis and associated complications are discussed in **Section 2.1.4**. Image courtesy of Dr. An Tang.

2.1.4 Cirrhosis

Cirrhosis represents the end-stage of many chronic liver diseases which cause necrosis of hepatocytes through a variety of insults (i.e. viruses, alcohol, fat). Pathologically, the characteristics findings in cirrhosis are: fibrosis, nodular regeneration and distortion of hepatic architecture (37). Eventually, functional liver tissue is replaced by non-functional scar tissue leading to liver failure.

Other complications of cirrhosis include: increased pressure in the venous system draining to the liver (*portal hypertension*) leading to bleeding from distended veins into the GI system (*variceal bleeding*), fluid accumulation in the abdominal cavity (*ascites*) and behavioral changes due to accumulation of toxic metabolites (*hepatic encephalopathy*) (29). Cirrhotic patients have a per-year risk of 1-8% for developing HCC (29). The clinical severity of cirrhosis is assessed by the Child-Pugh scoring system to advise clinical decisions regarding transplantation or hepatectomy.

Radiographic findings in advanced cirrhosis include hypertrophy of liver segments I, II and III with concurrent atrophy of segments VI and VII, likely related to alteration in hepatic blood flow (37). Frequent imaging features of cirrhosis seen on CT are outlined in **Figure 2.2**.

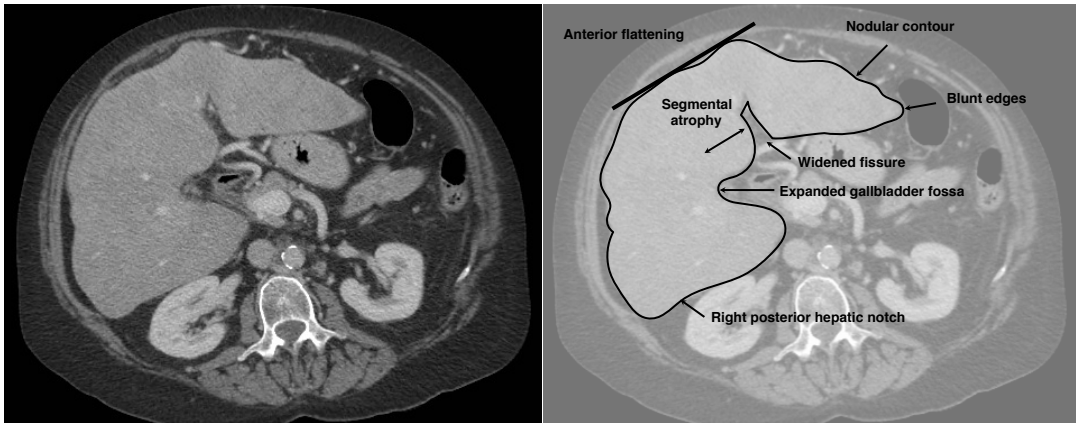

Figure 2.2: Frequent CT imaging features in cirrhosis
Imaging features in cirrhosis include: surface nodularity, widening of fissures and spaces (i.e. periportal, pericholecystic), atrophy of right anterior and left medial segments, hypertrophy of lateral segment, blunting of liver edges, posterior notching and anterolateral flattening. Other features include presence of regenerative nodules, siderotic nodules and secondary signs of portal hypertension. Image courtesy of Dr. An Tang.

2.1.5 Liver Cancer

Hepatocellular Carcinoma

Hepatocellular carcinoma represents the most common primary liver malignancy, constituting roughly 85% of all primary liver cancers (29). HCC typically develops in the context of cirrhosis. Major risk factors include: hepatitis B and C infection, alcoholism, biliary cirrhosis, food toxins, congenital biliary atresia, hemochromatosis, alpha-1 antitrypsin deficiency, type 1 glycogen storage disease and Wilson's disease (38). Both the incidence and associated mortality of

HCC have been increasing in Canada (29). This is partly attributed to more widespread hepatitis infections (39).

The Liver Imaging Reporting and Data System (LI-RADS) is an imaging classification system developed specifically for liver lesions (40). The LI-RADS score indicates the relative risk of a lesion being HCC in the context of predisposing risk factors, such as cirrhosis. Major imaging criteria for diagnosis of HCC include: arterial phase hyper-enhancement, portal venous or delayed venous phase washout, presence of capsule appearance and a specific pattern of threshold growth (40). An example of HCC with classic imaging findings is shown in **Figure 2.3**.

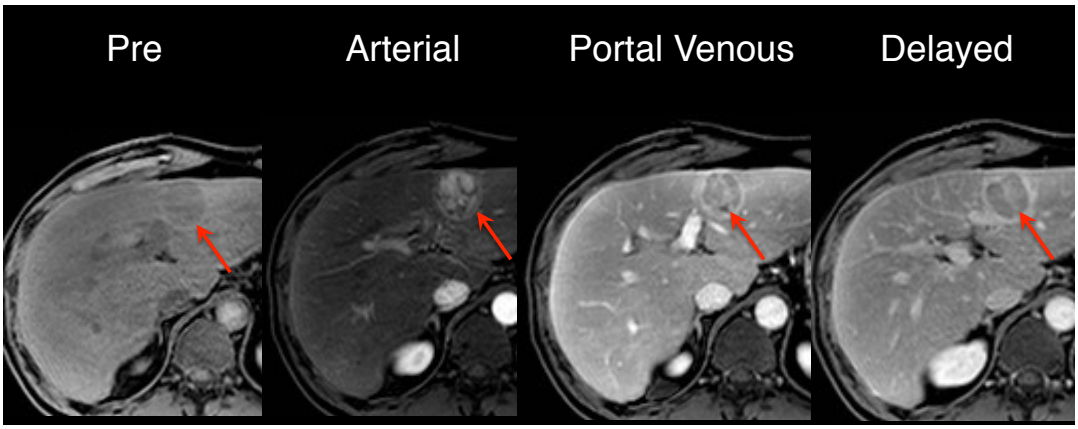

Figure 2.3: Classical imaging finding in HCC
A lesion highly suspicious for HCC is noted within the left lobe of the liver. The lesion displays characteristics imaging features of HCC: arterial phase hyperenhancement, contrast washout in portal venous and delayed phases, peripheral rim of smooth hyperenhancement in portal venous and delayed phases (capsule) and threshold growth (new lesion larger than 10mm represents threshold growth) (40). Image courtesy of Dr. An Tang.

Experts recommend screening for patients at high risk for developing HCC as cure rates associated with early diagnosis (prior to symptoms) approach 90% (29).

In North America, screening every six months with ultrasound examination is recommended for patients at sufficient risk (41). Common treatment options for HCC include partial hepatectomy, transplantation, ablation techniques, trans-arterial chemoembolization (TACE), chemotherapy, radiotherapy and palliative therapy.

Metastases

Metastases to the liver are very common, with studies suggesting that they occur 18-40 times more often than primary liver tumours (42). The most common sites that metastasize to the liver are: gastro-intestinal organs draining via the portal vein (i.e. colorectal, pancreatic, esophageal, gastric, neuroendocrine, gastrointestinal stromal), genitourinary (i.e. ovarian, renal, endometrial), breast, lung, melanomas and sarcomas (43).

Experience from colorectal carcinoma (CRC) can be used to demonstrate the burden of hepatic metastases. CRC represents the third most common cancer in the Western world (44). The liver represents the most common site of metastases with roughly 50% of patients with CRC developing hepatic metastases (44, 45). Hepatic metastases will be the main cause of mortality in two-third of patients with CRC (44).

2.2 Liver Volume as a Biomarker

Biomarkers

A biomarker is defined by the National Institutes of Health (NIH) as: "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (46). Various types of biomarkers and their characteristics are summarized in **Table II.I**. The characteristics of an ideal biomarker, regardless of intended purpose, are also described.

Table II.I: Types of biomarkers

References (46-48).

Biomarker	Characteristics
Antecedent	Used to identify the risk of developing a disease. Ex: Blood cholesterol concentration and risk of heart disease.
Screening	Used to screen for subclinical disease. Ex: Screening mammogram to detect early breast cancer.
Diagnosis	Used as a diagnostic tool to identify those with a disease. Ex: Blood glucose concentration for diagnosis of diabetes mellitus.
Staging	Used to stage or classify extent of disease. Ex: Prostate-specific antigen (PSA) concentration to reflect extent of metastatic disease.
Prognostic	Used as an indicator of disease prognosis. Ex: Tumour measurements after chemotherapy treatment.
Surrogate end-point	Used to substitute for a clinical endpoint, expected to predict clinical benefit. Ex: Human immunodeficiency virus (HIV) viral load as a surrogate end-point for acquired immunodeficiency syndrome (AIDS) diagnosis.
Ideal Biomarker	- Accurate - Reproducible - Sensitive and specific for given outcome - Easy to interpret - Acceptable to the patient - Data suggests that levels of biomarker will alter management

Liver Volume as a biomarker

The development of biomarkers for liver disease represents a growing research field within hepatology. The significant worldwide burden of liver disease, the late manifestations of symptoms with advanced disease, an intrusive reference test (liver biopsy) and the lack of validated tools to assess therapeutic efficacy are promoters of such research (49).

Liver volume determined from imaging examinations represents a non-invasive tool which has been explored in studies as a potential biomarker. For example, it is standard clinical practice to consider the future liver remnant (FLR) as a surrogate for hepatic reserve prior to hepatectomy (see **Section 2.3.1**).

Okazaki et al. assessed whether liver segment volume indexes calculated from MRI examinations varied in different forms of cirrhosis (50). They found that enlargement of the caudate lobe was more frequent in alcoholic cirrhosis than in virus-induced cirrhosis.

Zhou et al. explored the correlation between hepatic lobe volume variations in patients with virus-induced cirrhosis and severity of disease on 16-slice MDCT (51). They found that volume enlargement of the left lateral segment was absolute in Child-Pugh class A and B patients while enlargement of the caudate lobe was absolute in Child-Pugh class A patients.

Bora et al. elicited a positive correlation between hepatosteatosis and liver volume in patients with non-alcoholic fatty liver disease (52).

Crippin et al. investigated whether liver volume for ideal body weight could serve as a prognostic indicator in patients with cirrhosis (53). They found that liver volume could indeed predict survival in patients with cirrhosis caused by hepatocellular disease. Patients with smaller volumes had a statistically significant increase in transplant or death while those with larger volumes had a statistically significant survival advantage.

Imaging-based biomarkers

Imaging based techniques have recently been investigated as biomarkers of diffuse liver disease. For example, studies have attempted to quantify hepatic fat using MRI-based methods. These methods are non-invasive and may be as accurate and reproducible as liver biopsy, the current gold standard. Given the prevalence of NAFLD, early steatosis detection and measurement are crucial to institute appropriate management.

Liver biopsy is considered the gold standard for diagnosis of hepatic steatosis but has several limitations: it is invasive, has poor patient acceptance, has a risk of hospitalization of 1-5% and mortality rates between 0.01-0.1% (54-56). Further, it is prone to inter-observer variability and sampling errors (57). For these reasons, it is not considered acceptable for routine use in the fatty liver disease population. A reliable, reproducible and accurate method for fat quantification is thus needed.

Imaging-based techniques have emerged to assess liver fat content using magnetic resonance spectroscopy (MRS) (58, 59) and MRI (60-62). These techniques exploit the fact that fat resonates at a lower frequency than water when subjected to a homogeneous magnetic field (63). Such techniques can estimate the liver proton-density fat fraction (PDFF) which represents the fraction of proton density attributable to hepatic fat (60). PDFF maps can be translated pixel-by-pixel onto source images to generate parametric maps which illustrate the amount and distribution of fat throughout the liver (64).

Tang et al. recently introduced a novel volume-average biomarker: the total liver fat index (TLFI) in patients with NASH (65). It is defined as the product of the segmented liver volume by the average PDFF within the segmented volume. The study showed that a biomarker such as TLFI could be used to accurately monitor liver fat burden and its longitudinal change over time in the setting of a clinical trial.

Future studies may also incorporate liver volumetry to calculate other volume-averaged biomarkers of liver disease, such as,iron per unit volume (66). This

demonstrates the importance of liver volume as both a stand-alone and combined biomarker.

2.3 Clinical Indications for Liver Volumetry

As was demonstrated in the previous section, there are many reasons to perform either partial or complete liver volumetry. The established clinical indications for liver volumetry are: major hepatectomy, portal vein embolization and transplantation.

2.3.1 Future Liver Remnant (FLR) Prior to Major Hepatectomy

Liver resection, or hepatectomy, is performed for a variety of reasons including benign pathology (i.e. giant hemangiomas, hepatic adenomas, large cysts), malignant pathology (i.e. HCC, cholangiocarcinoma, metastases), infectious pathology (i.e. pyogenic or amebic abscess) and biliary or hepatic trauma. Hepatectomy is the treatment of choice for primary or metastatic liver tumours, providing the best chance for long-term patient survival (67, 68).

Major hepatectomy implies resection of four or more liver segments. The common types of major hepatectomy are demonstrated in **Figure 2.4**.

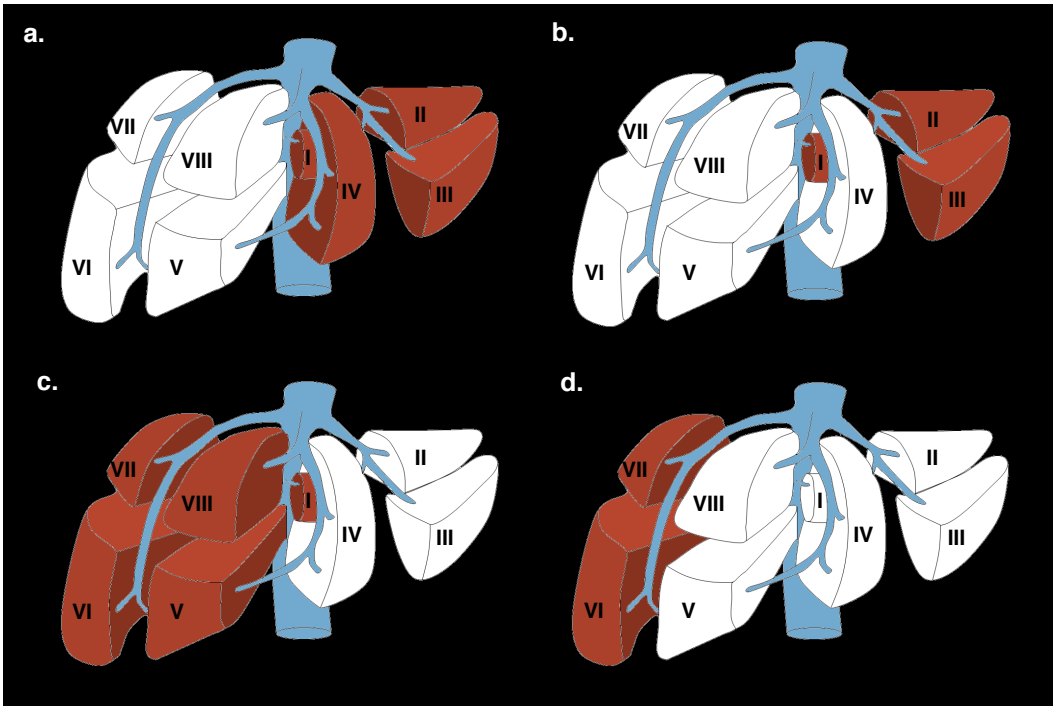

Figure 2.4: Types of major hepatectomy.
White segments are planned for surgical resection. (a) Complete right hepatectomy, (b) extended right hepatectomy, (c) complete left hepatectomy and (d) extended left hepatectomy. Extended right hepatectomy represents the most common type of major hepatectomy. Figure adapted from Terminology of liver anatomy and resections (8) (69).

The volume of the liver which remains post-hepatectomy is of particular clinical value. During surgical planning it is termed the "future liver remnant" (FLR) volume. The FLR volume post-hepatectomy is a direct indicator of residual liver function and post-operative outcome (70). It is also one of the only independent predictive factors of post-operative liver dysfunction (70).

There have been recent increases in extended hepatectomies as definitions of resectability have expanded, thus leaving less remnant liver (70, 71). Liver volumetry is currently indicated in patients undergoing major hepatectomy or having underlying liver disease (44, 70) to ensure adequate functional remaining liver.

It is vital to accurately establish both anticipated FLR volume and total liver volume (TLV) prior to hepatectomy. To be considered safely resectable, the FLR/TLV ratio must be > 26.5% for normal livers, > 40% in high-grade steatosis, and > 50% in cirrhosis, reflective of the underlying hepatic parenchymal quality (44, 72). This is visually demonstrated in **Figure 2.5**.

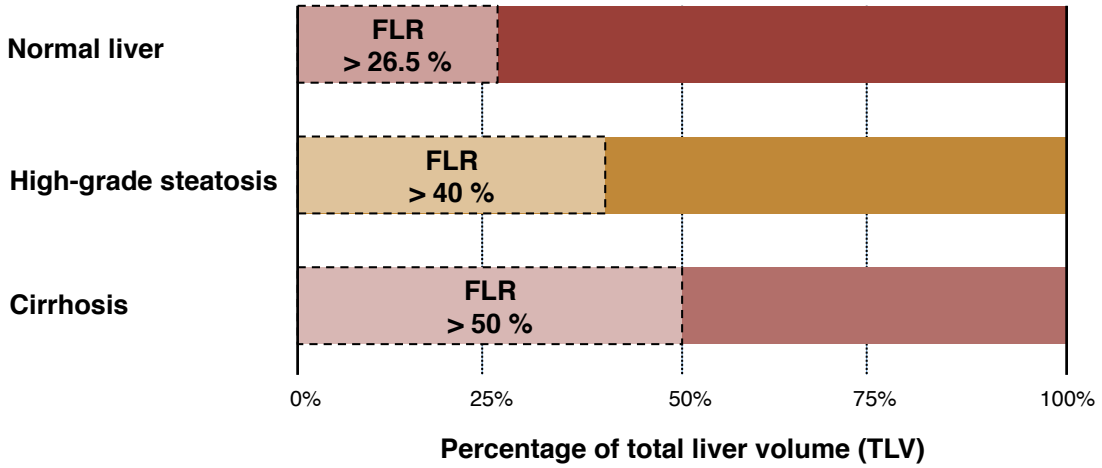

Figure 2.5: FLR/TLV ratio prior to hepatectomy.
To be considered safely resectable prior to hepatectomy, the FLR/TLV ratio must be > 26.5% in underlying normal livers, > 40% in high-grade steatotic livers and > 50% in cirrhotic livers (8).

Examples of FLR/TLV ratio calculations prior to hepatectomy and how they impact clinical judgment are demonstrated in **Figures 2.6** (normal), **2.7** (steatosis) and **2.8** (cirrhosis).

Normal liver

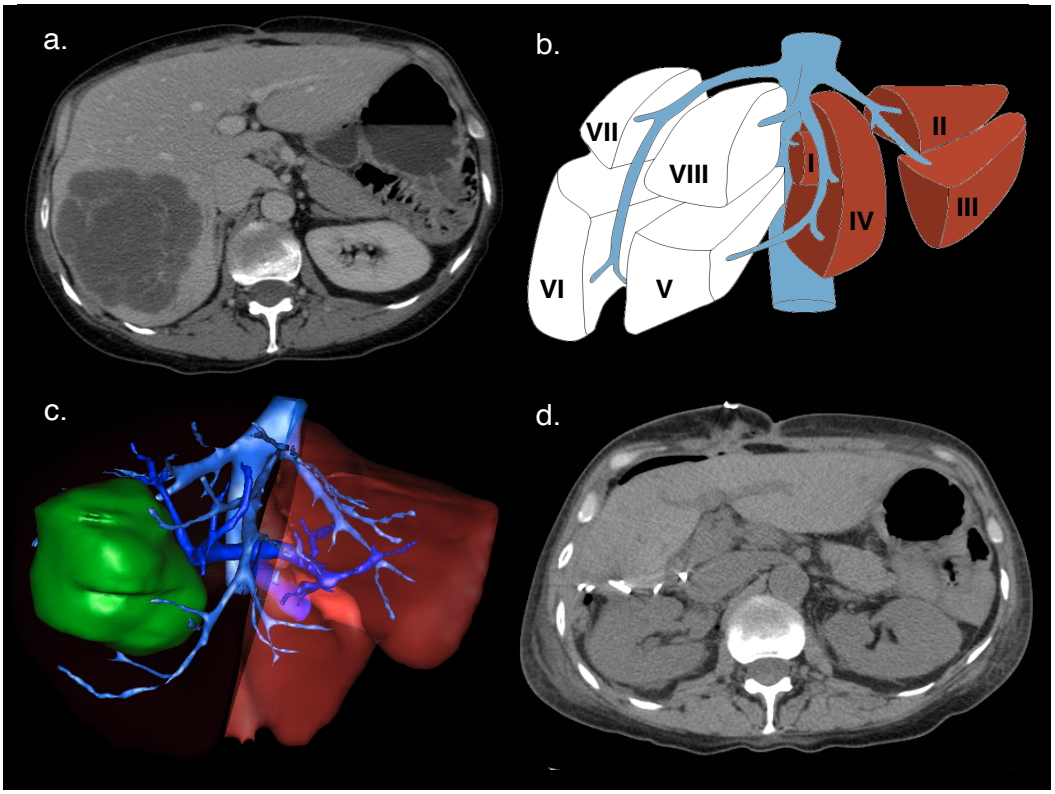

Figure 2.6: Future liver remnant volume calculation in normal liver.
(a) Axial enhanced CT image shows colorectal liver metastasis involving right posterior segments (VI and VII). (b) Resection diagram shows the intended complete right hepatectomy surgery planned. (c) 3D-rendering image shows surgical planning for complete right hepatectomy. FLR/TLV ratio was estimated to be 33%. (d) Axial unenhanced CT image of the same patient shortly after complete right hepatectomy. Actual FLR/TLV ratio was calculated to be 36% (8). Figure (c) courtesy of Dr. Franck Vandenbroucke-Menu.

Steatosis

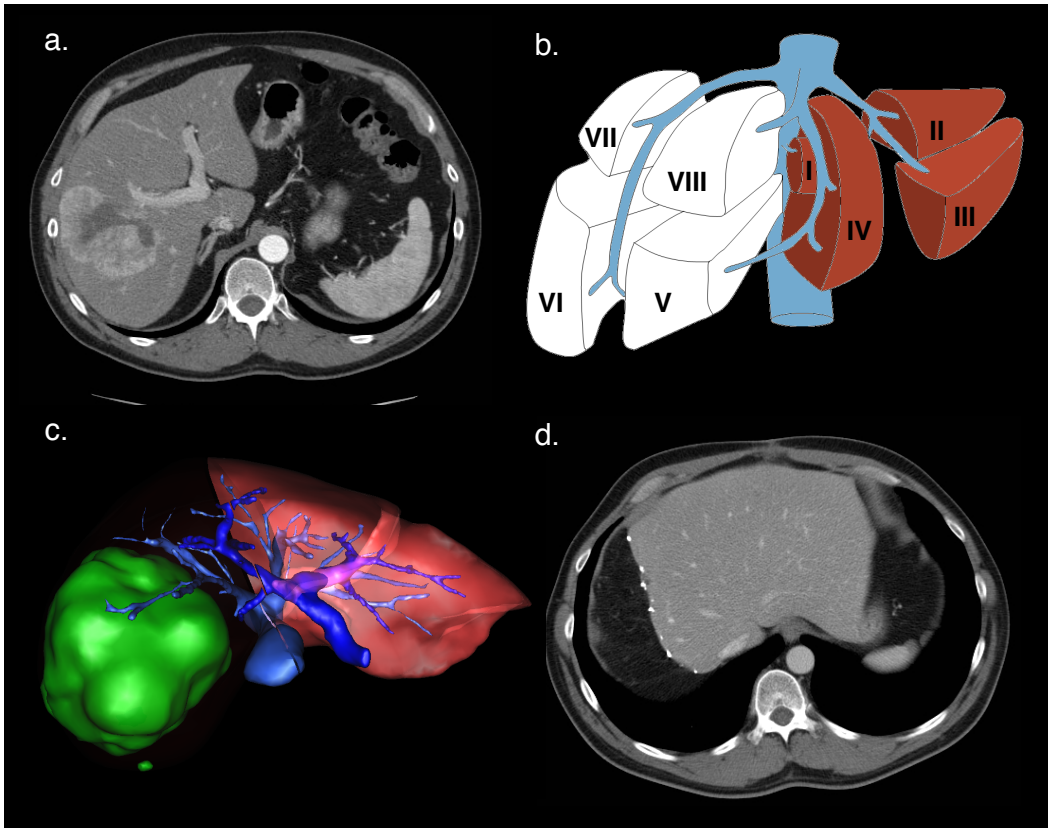

Figure 2.7: Future liver remnant volume calculation in fatty liver.
(a) Axial enhanced CT image shows colorectal liver metastasis involving segments V, VI, VII, and VIII. (b) Diagram showing the intended complete right hepatectomy surgery planned. (c) 3D-rendering image shows surgical planning for complete right hepatectomy. FLR/TLV ratio was estimated to be 46%. (d) Axial enhanced CT image of the same patient after complete right hepatectomy. Actual FLR/TLV ratio was calculated to be 60%. Figure (c) courtesy of Dr. Franck Vandenbroucke-Menu (8).

Cirrhosis

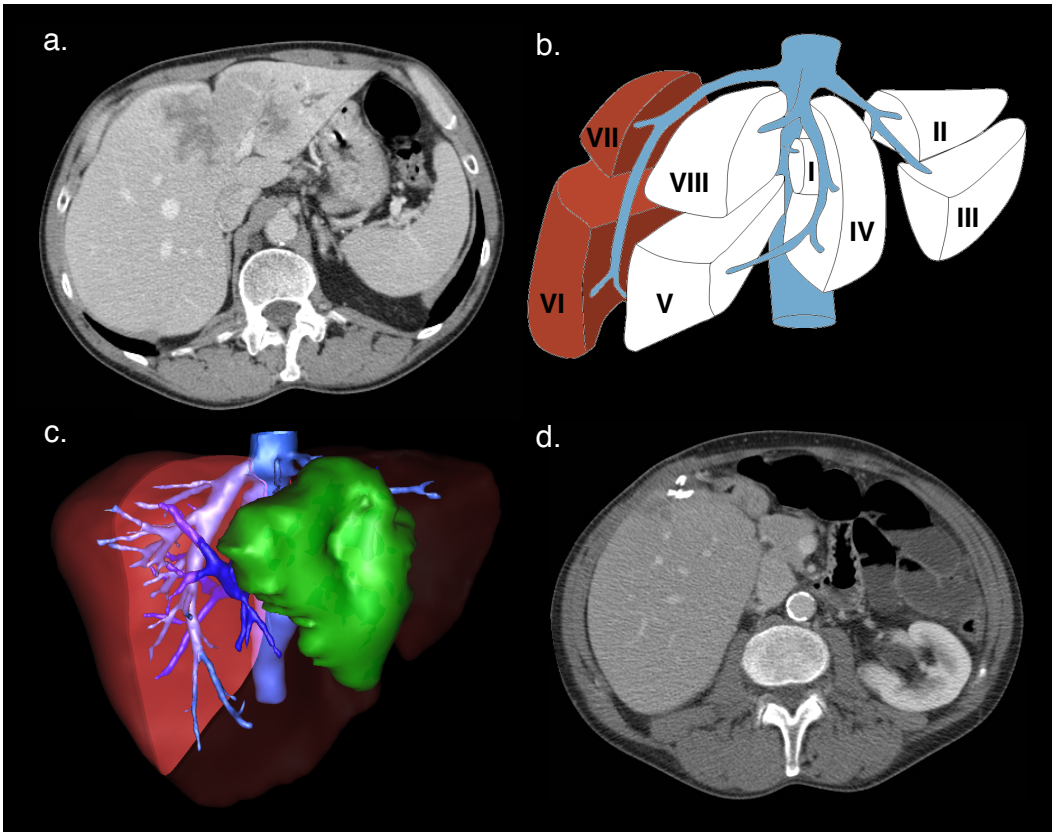

Figure 2.8: Future liver remnant volume calculation in cirrhotic liver.
(a) Axial enhanced CT image shows colorectal liver metastasis involving segments II, III, and IV. (b) Diagram showing the intended extended left hepatectomy surgery planned. (c) 3D-rendering image shows surgical planning for extended left hepatectomy. FLR/TLV ratio was estimated to be 45%. (d) Axial enhanced CT image of the same patient after extended left hepatectomy. Actual FLR/TLV ratio was calculated to be 49%. Figure (c) courtesy of Dr. Franck Vandenbroucke-Menu (8).

2.3.2 Portal Vein Embolization

Portal vein embolization (PVE) is a minimally invasive pre-operative procedure performed by an interventional radiologist. It entails the selective occlusion of the portal blood supply to certain liver segments, redistributing blood

flow towards segments which will remain post-hepatectomy. The ultimate goal is to reduce the risk of post-operative complications by intentionally causing hypertrophy if the residual liver (i.e. increase mass of the anticipated FLR). Studies have shown improvement in liver function post extended hepatectomy in patients undergoing PVE compared to without (44) (73).

Indications for PVE rely on factors which may impact the FLR volume required for adequate post-hepatectomy liver function (44). Underlying liver disease, recent chemotherapy and the extent of the planned resection are all important factors. PVE is indicated when the FLR/TLV ratio is ≤ 20% in a normal liver, ≤ 30% in the setting of recent chemotherapy, or ≤ 40% in a fibrotic or cirrhotic liver. This is visually demonstrated in **Figure 2.9**.

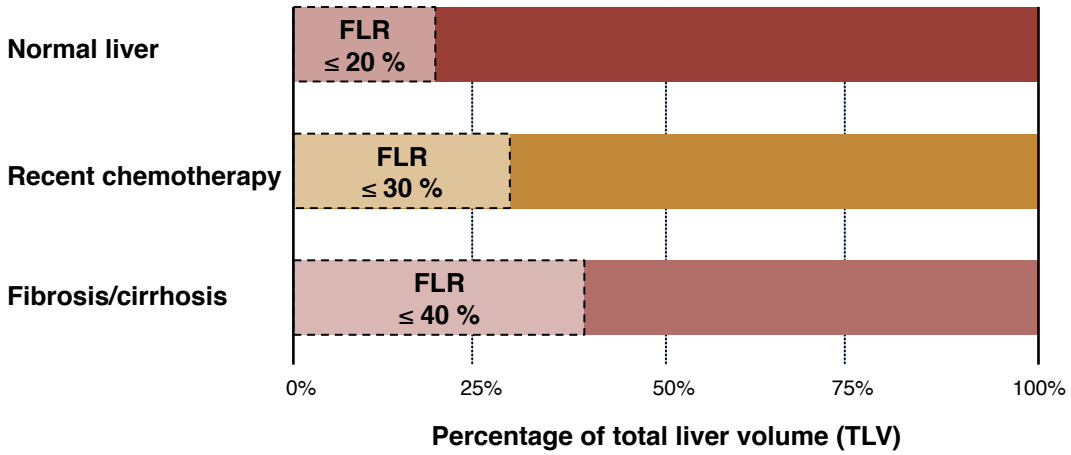

Figure 2.9: FLR/TLV ratio prior to portal vein embolization.
Portal vein embolization is indicated when the FLR/TLV ratio is ≤ 20% in an underlying normal liver, ≤ 30% in the setting of recent chemotherapy, or ≤ 40% in a fibrotic or cirrhotic liver (8).

Liver volumetry is clinically indicated to initially calculate the FLR/TLV ratio and 3-4 weeks after PVE to assess volume and extent of hypertrophy (44). An example of a case requiring PVE prior to right hepatectomy is shown in **Figure 2.10**.

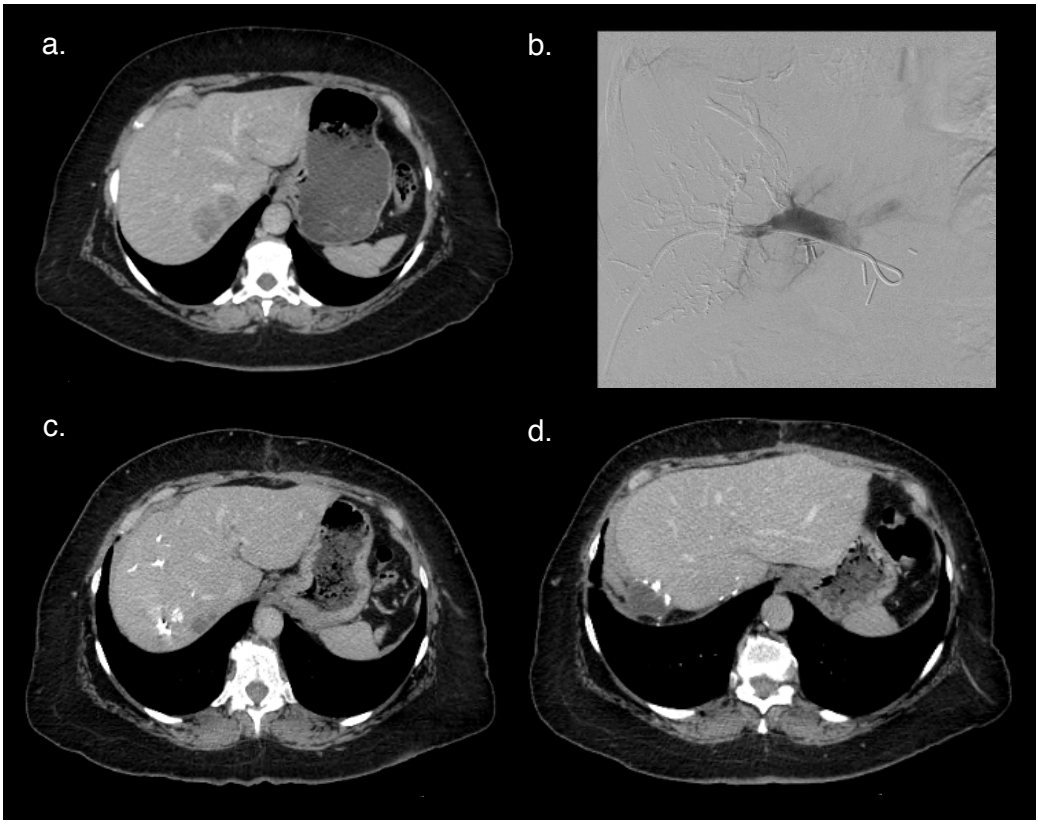

Figure 2.10: Portal vein embolization prior to right hepatectomy
(a) Axial enhanced CT image shows colorectal liver metastasis involving segments V, VI, VII (only VII shown). (b) Embolization of the portal vein branches to segments V through VIII was performed using a Lipiodol-glue mixture. Final portogram is shown. (c) Axial enhanced CT image obtained 1 month after right PVE shows hypertrophy of future liver remnant. (d) Axial enhanced CT image of the same patient after right hepatectomy (8).

2.3.3 Living Donor Liver Transplantation

Due to increasing demand and scarcity of cadaveric livers, alternatives have been sought to basic orthotopic liver transplantation. Transplantation of the left lateral segment from a living donor is performed for the pediatric population, but this does not provide adequate hepatic volume for adult recipients (74). Similarly,

cadaveric split-liver transplantation may not provide adequate hepatic volumes for two adult recipients (75).

Living donor liver transplantation is being increasingly performed, exploiting the regenerative capacity of the liver. Pre-operative imaging of the donor is performed to exclude hepatic lesions, assess for diffuse liver disease and assess vascular and biliary anatomy (24). Moreover, pre-transplant liver volumetry is indicated as appropriate graft size is a major indicator of successful clinical outcome for both donor and recipient.

In living donor transplantation, a FLR-TLV ratio of 30-40% is required by the donor for survival (76, 77). In the recipient, the graft size to recipient body weight ratio ideally must be higher than 0.8-1.0% (78). Alternatively, the graft size to standard liver volume (calculated from body surface area) ratio must be higher than 50% for the recipient (79).

Insufficient graft size may lead to "small-for-size syndrome" in the recipient. In this syndrome the graft is too small to meet functional demands resulting in liver failure and possibly death in the absence of re-transplantation (80). An example of size incompatibility during living donor liver transplantation is provided in **Figure 2.11**.

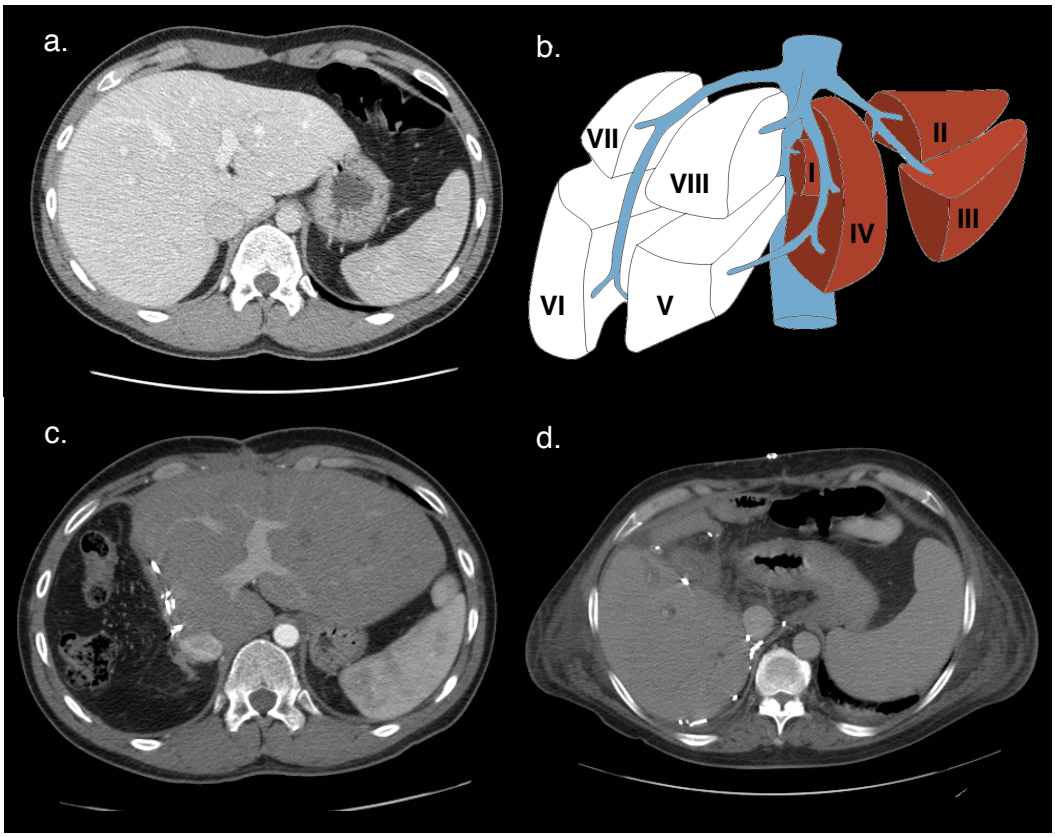

Figure 2.11: Size incompatibility in living donor liver transplantation.
In this live liver donation, both the donor and recipient had transient hepatic insufficiency due to small-for-size liver: (a) Axial enhanced CT image of a 26-year-old living liver donor. The total liver volume (TLV) was 1754 mL. The donated liver volume was 980 mL and the residual liver volume was 774 mL (44.2% of the TLV). (b) Diagram showing the intended right split liver surgery planned for living donor liver transplantation. (c) Post-liver transplantation axial enhanced-CT image showing hypertrophied left liver of the donor. (d) Post-liver transplantation axial enhanced-CT image of a 53-year-old man who was the recipient of the right liver transplant (8).

2.4 Reference Methods for Liver Volumetry

2.4.1 Formula-based

In 2002, Vauthey et al. established formulas to calculate total liver volume based on body surface area and body weight (73). Total liver volumes were measured in 292 patients from segmentation of CT scan images at four different sites. The study assumed that the volumetric methods used to calculate total liver volumes correlated with actual liver volume. Formulas were established using patient's body surface area (BSA) and body weight using regression analysis as follows:

1. TLV = - 794.41 + 1267.28 x BSA (square meters)
2. TLV = 191.80 + 18.51 x weight (kilograms)

The authors cautioned that these formulas should be considered estimates due to the variability in correlation of total liver volume with body surface area (r^2 = 0.46) and body weight (r^2 = 0.49). The study had excluded patients with any kind of diffuse liver disease thus limiting extrapolation to these patient groups. Though other formulas to estimate total liver volume have been proposed, the ones described in this study remain widely used in the clinical setting.

2.4.2 Surgical Specimen

In the validation of their automated liver volumetry methods, many authors have used surgical resection specimens as the volumetric reference standard.

Hermoye et al. compared the accuracy and repeatability of MRI-based semiautomatic liver volumetry with surgical graft volume in living liver transplant donors (79). The donors underwent either left lateral segmentectomy, complete left or complete right hepatectomy. The liver grafts were flushed successively through a portal vein cannula. The grafts were then weighed using a calibrated scale. Graft

weight was converted to graft volume by assuming a liver density of 1.0 g/cm^3, and this was used as the reference standard.

Nakayama et al. compared in vivo CT-based automated volumetry with surgical liver volume using native livers of patients awaiting living related liver transplants (14). First, the relationship between liver weight and liver volume was determined by means of a regression analysis in seven other transplant patients. Liver volume was determined by placing excised livers in a water bath filled with distilled water at 25°C and measuring water displacement. Liver weight was determined by directly weighing the specimens on a scale. This provided a regression line (y=1.06) for the relationship between liver weight and volume.

Thirty-five recipients had their native livers excised prior to obtaining a liver related transplant. The gallbladder, portal structures, attachment ligaments and tissues were removed and all blood was drained. The specimens were then directly weighed and their volume was determined based on the previously created regression line. The study found excellent correlation between liver weight and volume ($r=0.957$, $p<0.01$). This volume served as the reference standard in the comparison with the automated method.

Lemke et al. developed equations to calculate the expected intra-operative weight and volume of living donor's right liver lobes using pre-operative CT-based volumetry (15). The right liver lobe specimens were flushed through all vascular structures at the liver hilum immediately after resection. Specimens were weighed with electronic laboratory scales. The specimens were then placed in a cylindrical 6L glass container filled with sterile 4°C physiologic saline. The volume of the displaced liquid was measured and was assumed to correspond to volume of the immersed liver specimen. Regression analysis determined two linear equations which can be used to calculate intra-operative volume ($V_{intraop}=(0.656 \times V_{preop})+87.629mL$) and weight ($W_{intraop}= (0.678\ g/mL \times V_{preop})+143.704g$) from pre-operative CT liver volume. The study found adequate correlation between preoperative volumetric measurements and intraoperative volume ($r=0.834$, $p<0.001$) and weight ($r=0.870$, $p<0.001$).

However, the study found a substantial discrepancy (mean deviation 34.3%) between pre-operative and intra-operative hepatic volume measurements. Possible explanations included: perioperative loss of blood (81), deviation between assumed physical density of the liver ($1 g/cm^3$) and actual physical density (81) and influence of altered perfusion states (secondary to hemodynamic drugs, blood loss, clamping of vessels) on liver volume (15, 82).

2.4.3 Manual Segmentation

The most common current method to estimate the liver volume involves manually delineating the liver outline on consecutive CT and MRI images, a process called "segmentation". Manual segmentation has been used in many studies as the reference standard for evaluating CT-based volumetry (13, 14, 17, 18, 27) and MR-based volumetry (83-90).

Typically, axial CT and MRI images are saved as Digital Imaging and Communications in Medicine (DICOM) files and uploaded to imaging post-processing display software. Image analysts then manually contour the liver using a cursor. The most basic tool is a pencil or spline widget which positions nodes along the liver boundary **(Figure 2.12a)**. The aim is to delineate the liver outline on each axial slice. Large vessels abutting the liver periphery such as the main portal vein and inferior vena cava are usually excluded, while vessels completely surrounded by liver parenchyma are included. The number of pixels within each contour provides a cross-sectional liver area on a slice-by-slice basis **(Figure 2.12b)**. This area is then multiplied by the slice thickness and the summation of each section volume provides the total liver volume for each patient. There is significant variability in terms of imaging post-processing software and type of contouring tools used.

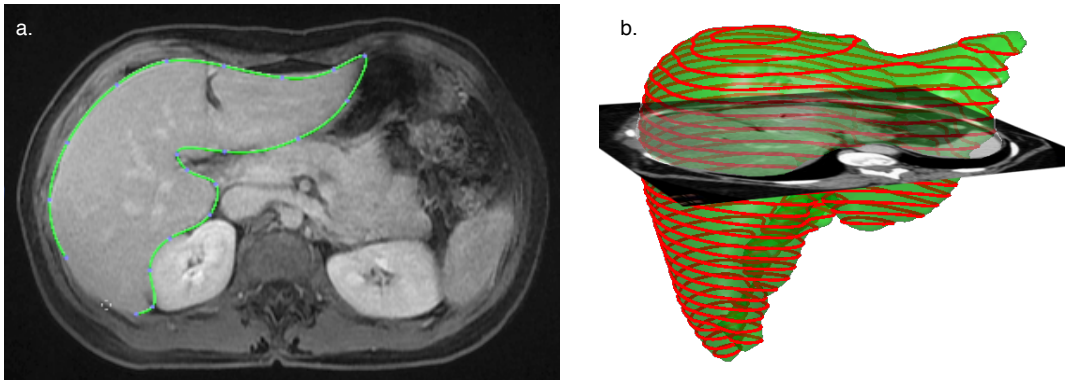

Figure 2.12: Manual segmentation.
(a) Manual liver segmentation consists of contouring, or tagging, pixels belonging to the liver on every slice of a CT or MRI image. The most basic tool is a pencil or spline widget where nodes are manually positioned along the liver boundary. Image obtained using Osirix image post-processing software (Osirix Foundation, Geneva, Switzerland). (b) Once the segmentation complete, volumetry is obtained based on pixel size and slice spacing (8).

Manual liver volumetry has been criticized for being overly time-consuming for clinical purposes, often requiring 30-90 minutes to assess the liver volume of one patient (26). Further, there is inherent inter-observer and intra-observer variability given the visual judgment required when tracing contours. Precision in manual segmentation is dependant on various factors such as user experience, sharpness of liver boundaries, the window level settings affecting image display, computer monitor settings and user vision characteristics (28), all of which can introduce variability. Despite these shortcomings, manual segmentation remains the most popular reference standard for the validation of CT and MRI-based liver volumetry methods.

2.5 Automated Liver Segmentation

The development and validation of automated liver segmentation methods represents a very active research area. Various liver segmentation pitfalls (see

Section 3.1) make it difficult to design an automated tool that is functional in every clinical situation. Therefore, clinicians often prefer manual segmentation as it is easier to implement though cumbersome and not adapted to clinical reality.

This section addresses various automated segmentation methods (contour optimization, interactive and fully automated) in order of increasing complexity. A outline is provided in **Figure 2.13**.

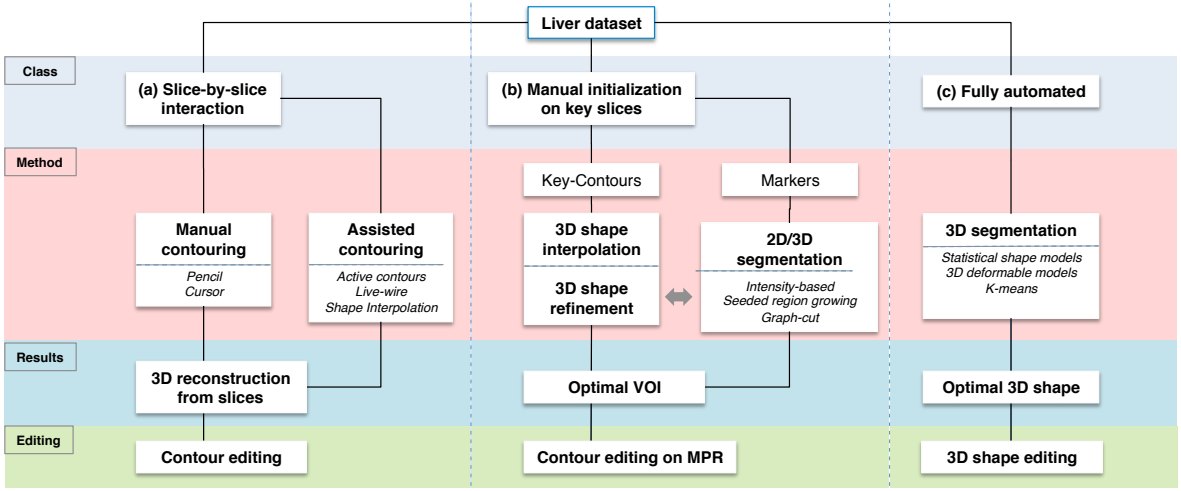

Figure 2.13: Automated liver segmentation outline.
Liver segmentation methods may rely on (a) manual, (b) semiautomated, or (c) fully automated workflows. Most workflows require a combination of 2D or 3D initialization, refinement, and editing techniques. VOI = Volume of interest. MPR = Multi-planar reconstruction (8).

2.5.1 Contour Optimization Techniques

As described in **Section 2.4.3**, manual segmentation consists of contouring the liver on each slice of a CT or MRI exam. This is a very time-consuming procedure and precision is dependent on a variety of user-dependant factors. To improve time efficiency and precision, several automated contour optimizations techniques have been proposed which can supplement manual segmentation.

Active contours

In the active contour approach (91), the segmentation is defined by an outlining contour for which internal forces (rigidity), counterbalance external forces, defined by the underlying image data. For this iterative algorithm to reach convergence on the desired segmentation, parameters need to be carefully tuned and the image adequately pre-processed. This equilibrium can be difficult to achieve when the initial solution is too far from the liver boundary, which causes the active contour to fall into a local minimum or leak into adjacent organs. Implemented methods therefore use this algorithm in combination with a robust initialization method that can provide an as close as possible solution so that the contours converge in few iterations **(Figure 2.14)**.

Of note, the active contours technique serves as the basis for software created by Tomovision SliceOmatic® for manual segmentation. This software was used to establish the manual reference standard in the studies presented in **Sections 4** and **5**.

Figure 2.14: Active contours technique.
(a) For a given axial slice, an image analyst roughly contours the liver using a cursor. (b) and (c) These contours (snakes) then evolve a coarse contour based on image salient features and "snap" to the liver contour (8).

Livewire

In the livewire approach (92), an image is interpreted as a weighted graph. Image pixels are represented by graph vertices and graph edges connect adjacent

pixels with their weights representing the cost of connection (23). As the user clicks on the boundary to establish a "seed point", the possible minimal cost paths to all other points on the image are computed. Another boundary point can then be chosen via the "free point" (the mouse's current position). As the mouse is moved over other points, the boundary behaves like a livewire, connecting the seed point with the free point via a minimal cost path along the liver edge **(Figure 2.15)** (23). The 2D livewire technique serves as the basis for software created by MeVis HepaVision® for manual segmentation (23).

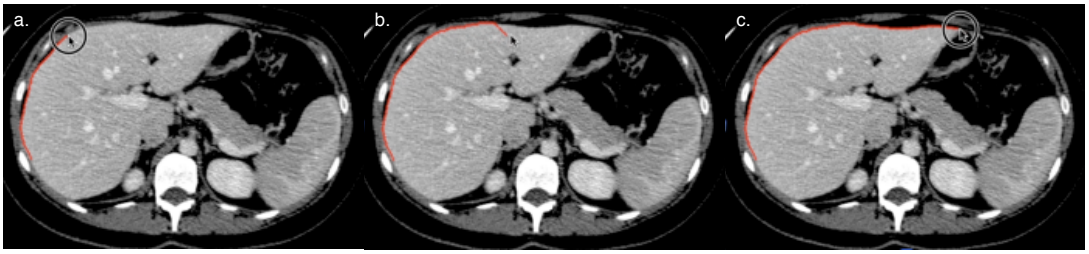

Figure 2.15: Livewire technique
(a) The user sets the "seed point" by clicking on the liver boundary. (b) As the mouse is moved over other points, the boundary behaves like a live wire connecting the seed point to the free point (c) via a minimal cost path along the liver edge (8).

Shape Interpolation

Shape interpolation allows the user to obtain a plausible complete 3D shape from a limited number of contours (93). This technique will be discussed in detail in **Section 3.2**.

2.5.2 Interactive Segmentation Techniques

Interactive or semiautomated segmentation techniques aim to reduce the amount of user input by relying on various types of interactions to steer the

segmentation process. These methods are most intuitive in 2D, but also supports 3D and 4D segmentation.

Often, the user may have to initialize the segmentation process by positioning nodes manually around the liver contour.

Intensity-based methods

Among intensity-based methods, traditional approaches such as region-growing (94) **(Figure 2.16)** have been used, which perform voxel or texture classification in feature space. These methods are usually computationally efficient and produce excellent results when the liver's intensity is homogeneous. The main downside is that no shape control is enforced, leading to rough edges and important leakage across the organ's boundary. Given that the heart, the stomach, the spleen and intercostal muscles may have similar densities, no interface clearly defines the liver. Consequently, these adjacent organs are often partly included and hard to dismiss automatically, requiring substantial user interaction to achieve desired results.

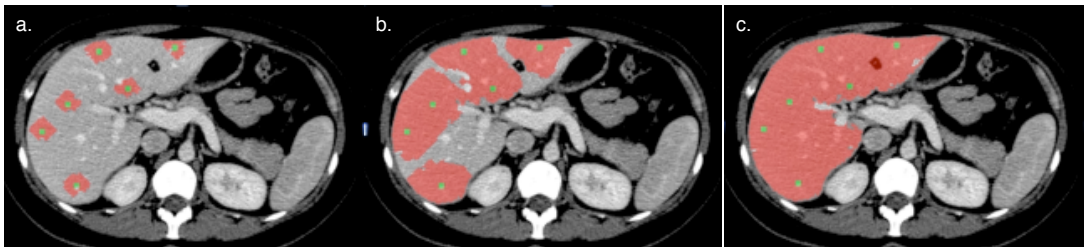

Figure 2.16: Seeded region-growing technique
(a) Seeds are initially positioned inside the regions of interest. (b) and (c) Pixels are iteratively aggregated if their intensity is similar to those already tagged (94) (8).

Graph-cut

The graph-cut requires the user to roughly paint some foreground an background pixels (95). Based on graph analysis and optimization, a cut is then performed to separate the foreground and background areas in the most homogeneous regions.

2.5.3 Automated Segmentation Techniques

Other methods aim at complete automation of liver segmentation. Automated segmentation methods usually perform well on typical datasets but often require manual adjustments on pathological and unusual cases.

Statistical shape models

To further constrain segmentation outcomes, global shape models were proposed to prevent deviation from a reasonable liver shape and to prevent segmentation leakage. Previous studies have used a single prior model which was subject to local curvature and global shape forces (96). Since the liver varies considerably amongst patients, a single prior shape is limited in coverage of the wide range of liver morphologies. Therefore, most recent approaches depend on statistical shape models (SSM) to expand the range of admissible liver shapes (97). In these approaches, the image information drives the deformation of a surface mesh parameterized by its principal shape components within a hyperspace of admissible shapes **(Figure 2.17)**.

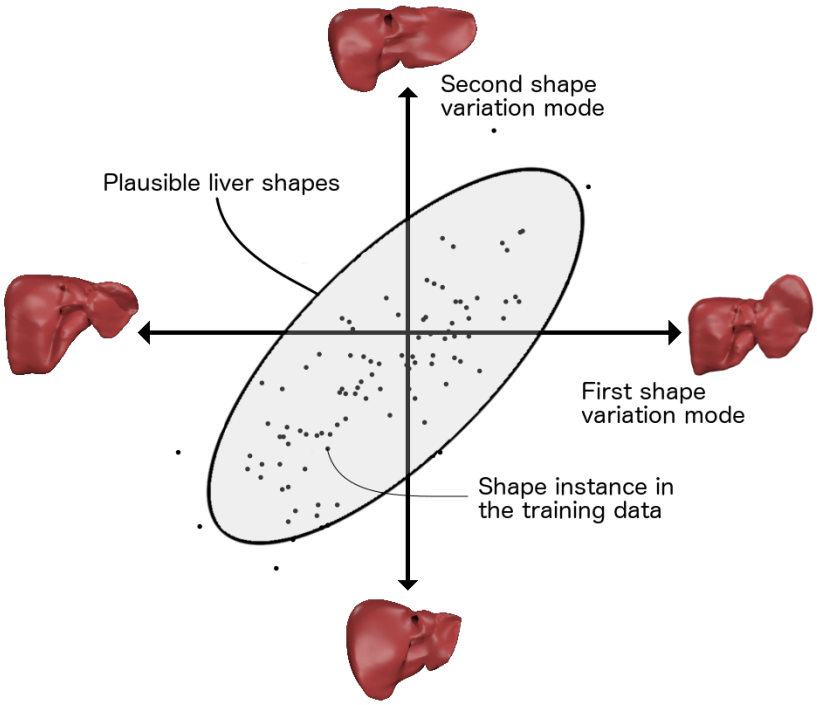

Figure 2.17: Statistical shape models

To restrict the segmentation to a set of admissible liver shapes, a shape database is compiled, from which any new liver shape is expressed by a set of parameters called modes of variation. The various modes of variation (typically roughly 30 modes) are adjusted to fit the liver shape on image features. Statistical shape models impose hard restriction on the segmentation outcome by integrating prior shape. However, training data cannot capture all variation and therefore are sometimes too limiting to accurately model specific livers (8).

The SSM approach provides a clear advantage over traditional active contours. Further, it has been shown to reliably produce accurate segmentation in noisy data and remain robust to common segmentation pitfalls such as contact with adjacent organs by imposing important shape constraints (27). However, the training phase involved is elaborate and represents a disadvantage. Moreover, the resulting model might be too constraining to reach liver shapes with features that weren't considered in the training data, leading to potential limitations on pathological or previously resected livers.

Recently, a liver segmentation competition compared liver segmentation algorithms on a common database of contrast-enhanced CT images at the MICCAI 2007 Grand Challenge (27). The study supported the use of both statistical shape information and model-based approaches to accurately represent liver structure variability (27). The authors also found that while statistical-based methods are certainly beneficial for automated segmentation, they are too constraining to reach precision and need to be coupled with a secondary segmentation method, such as graph-cut.

3D deformable models

A 3D extension of the active contours approach **(Section 2.5.1)** led to the 3D deformable models technique. This technique results in a 3D surface mesh which evolves inside the patient's dataset (26). The mesh may be subject to a non-rigid registration scheme based on Laplacian mesh deformation (98) which precisely deforms the shape towards the liver boundary. If carefully tuned, an approximate shape or a simple sphere is sufficient to initialize the process. This technique is described in detail in **Section 3.2.2**.

k-means

Classification algorithms, such as the k-means, attempt to label pixels as liver parenchyma based on their intensity and their texture properties (99). Since this often leads to coarse segmentation, it is generally combined with morphological operations such as mask erosion or dilation to eliminate structures having the same intensity.

2.5.4 Advanced Segmentation Methods

Combination methods

The presented segmentation methods are seldom used on their own. Advanced segmentation strategies often rely on combinations of various concepts. For instance, pixels classified by k-means with high confidence can be used as a background/foreground labeling for a graph-cut optimization. Statistical shape models may be used as robust initialization for unconstrained 3D active contour models.

A summary of advantages and limitations of various segmentation methods illustrated in the previous sections is provided in **Table II.II**. Green boxes indicate desirable features, whereas red boxes indicate limitations of various strategies. Manual segmentation methods require significant interaction and are time-consuming. Assisted contouring improves reproducibility and robustness. Semi-automation shortens interaction time. Fully automated segmentation methods are reproducible and rapid, but complex to implement and may fail.

Table II.II: Summary of advantages and limitations of various segmentation methods.
Reference (8)

Approaches	Methods	Reproducibility	Robustness	Time	Interactivity	Complexity of implementation
2D	Manual	↑	↑	↑ ↑	↑ ↑	↓ ↓
	Manual with assisted contouring	↑ ↑	↑ ↑	↑	↑	↓
2D/3D	User-initialized & semiautomated	↑	↓	↓	↓	↑
	Fully automated segmentation	↑ ↑	↓ ↓	↓ ↓	↓ ↓	↑ ↑

Outsourcing

Alternatively, private companies such as MeVis Medical Solutions® and EDDA Technology, Inc. offer a variety of segmentation, visualization, and analysis tools using their own proprietary software.

As noted from the MeVis website (http://www.mevis.de/en/): "On the basis of computed tomography (CT) or magnetic resonance imaging (MRI) data, the MeVis Distant Services (MDS) team of specialists develops detailed treatment scenarios, including three-dimensional presentations of the liver anatomy, exact volume quantifications and risk analyses."

Typically select DICOM images are sent to the company's file transfer protocol server. Segmentation images and mesh are then returned to the customer. The mesh can be manipulated a posteriori for surgical planning.

MRI-based methods

Though liver segmentation literature mostly pertains to CT imaging, there has been limited research more specifically oriented toward MRI-based methods. MRI offers a broader range of tissue contrast, providing additional tissue characterization sequences for detection of liver pathologies. Being less invasive than CT, it is suggested that it could satisfy preoperative imaging needs for volume measurements and surgery planning simultaneously.

MRI segmentation involves additional challenges as described in **Section 3.1.3**, including: poor resolution, motion artifacts due to long acquisition time and most notably, intensity inhomogeneities. Since there is a wider range of acquisition parameters for MRI and no common segmentation database is yet publicly available, the comparison and validation of segmentation algorithms are still subject to database variability. Few MRI-based automated or semiautomated segmentation techniques have been reported in the medical literature. **Table II.III** summarizes a selection of MRI-based techniques which have recently been

reported in the literature. Key results from each are discussed elsewhere in the manuscript.

Table II.III: Summary of MRI-based liver segmentation techniques

Authors	Patients	Segmentation Techniques	MRI Sequence	Reference Standard
Mazonakis et al. 2002. (83)	n=38 (consecutive)	**Manual-Semiautomated.** Stereology (Point-counting)	2D FLASH T1 gradient echo	Manual
Hermoye et al. 2005. (79)	n = 18 (living transplant donors)	**Semiautomated.** Geometric deformable models, level-set	Contrast-enhanced T1, fast-field echo	Surgical graft weight
Farraher et al. 2005. (85)	n=27 (normal and pathological livers)	**Semiautomated.** Dual space clustering algorithm	Mixed fast spin-echo pulse	Manual
Sahin et al. 2006. (100)	n=5 (cadaveric livers)	**Manual-Semiautomated.** Stereology (Point-counting)	Spin-echo	Manual
Siewart et al. 2010. (101)	n=10 (routine examinations)	**Automated.** Intensity-based, morphological region-growing	3D T1 LAVA	Manual
Gloger et al. 2010. (102)	n=20 (variable)	**Automated.** Probability map, region-growing	T1 FLASH 3D VIBE	Manual
Rusko and Bekes. 2011. (103)	n=60 (variable)	**Automated.** Probabilistic models, intensity-based	LAVA	Manual
Torkzad et al. 2012. (88)	n=41 (consecutive)	**Manual-Semiautomated.** Stereology (Point-counting)	T1	Manual
Huynh et al. 2014. (90)	n=23 (variable liver disease)	**Automated.** Geodesic active contours, level-set	T1 LAVA	Manual

Lopez-Mir et al. 2014. (104)	n=17 (variable training dataset)	**Automated.** watershed transform and stochastic partitions	Contrast-enhanced T1	Manual

A novel segmentation method

In **Section 3** we propose a novel method that can be effectively used for liver segmentation for both CT an MRI images, which has rarely been previously reported in the medical literature. Specifically, we address the shape initialization problem with implicit shape modeling and combine it with a shape deformation scheme based on Laplacian mesh optimization. It is independent of training data, requires modest user interactions and is robust to a wide variety of pathological cases. Two correction tools based on the same deformation scheme are further implemented which allows the user to further improve the segmentation.

3 Segmentation Software

3.1 Software Concept

Based on our combined experience in liver imaging and software development, our research team developed a list of features expected from an automated liver segmentation solution. These needs are summarized in **Table III.I** and expanded on in further detail below.

Table III.I: Features expected from an automated liver segmentation solution

Categories	Needs
Clinical needs	-Efficient clinical workflow -Possibility to delegate segmentation to technologist/image analyst -Ability to edit segmentation results -Trackability -User-friendly interface -Accurate -Reproducible -Robust -Fully automated vs. semiautomated -Efficient (<10 minutes/case) -Compatible with CT and MRI images
Quantitative output	-Uniform segmentation approaches -Can report accurate whole liver volume (< 5% error) -Can extract 3-dimensional liver mesh
Minimal Error	-Minimize: -Error linked to method initialization -Error linked to modality -Error linked to patient anatomy

3.1.1 Clinical Needs

Liver segmentation solutions are developed to address the clinical needs of physicians and to implement research and development methods developed by biomedical engineers. However, it is unrealistic to expect users in these domains to

perform liver segmentation on a daily basis. Current practice includes delegation of segmentation tasks to radiology technologists, image analysts or trainees. Being provided with the vital information obtained from liver segmentation allows physicians to concentrate on patient care, thus optimizing clinical workflow.

A liver segmentation graphical interface should be intuitive enough to allow routine operation by users with variable levels of training. Thus, key anatomical landmarks and image projections required for liver segmentation should be clearly identifiable. Furthermore, the software should allow for quality control (i.e. validation and correction) of the segmentation results by the clinician at a later time, if needed.

The graphical interface employed for our segmentation method included axial, coronal and sagittal planes of the abdomen; similar to an imaging workstation during routine clinical practice. A 3D projective view where all initial drawn contours could be visualized was also available. An example of the graphical interface used for semiautomated segmentation is provided in **Figure 3.1**.

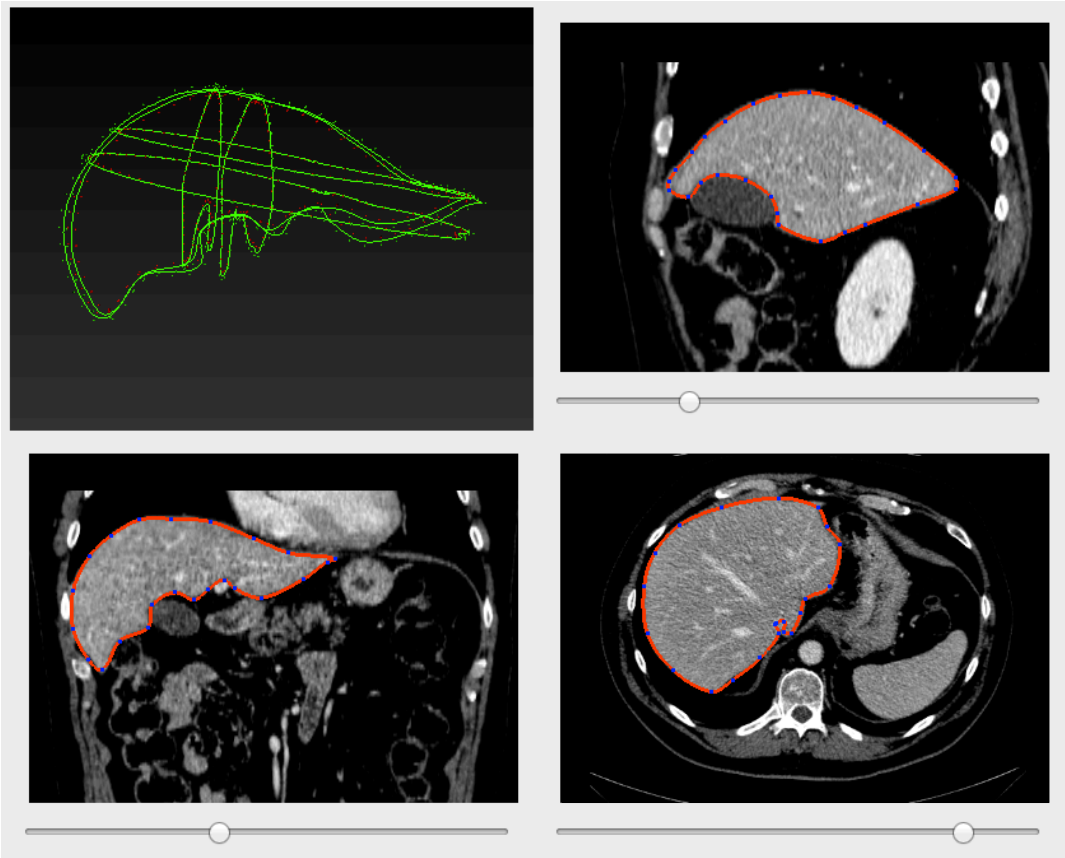

Figure 3.1: Graphical interface for semiautomated liver segmentation.
The user is provided with axial, coronal and sagittal planes of the abdomen as well as a 3D projective view where all initial drawn contours can be visualized. The user can scroll through images similar to an imaging workstation interface.

The accuracy, precision, and robustness of the automated segmentation method should be validated against a reference standard. Common reference standards used in validation of liver volumetric methods are described in **Section 2.4**. For our validation purposes we chose a manual segmentation method supplemented with an active contours technique. Our validation scheme is further described in **Sections 3.4**, **4** and **5**.

While full automation may appear ideal, it is often associated with additional problems. Heimann et al. describe fully automated liver segmentation methods as those where "each algorithm had to use the same set of parameters for all test

images" (27). Conversely, interactive or semiautomated segmentation methods are those which "require a certain amount of user interaction to complete" (27). This user interaction can range from placing a single seed point within a region on interest to extensive manual alterations of the segmentation mesh.

As shown in **Table II.II**, though fully automated segmentation methods are reproducible and rapid, they are also complex to implement and lack robustness. Inevitably, automated methods will generate errors due to unforeseeable anatomical variants or technical challenges. The MICCAI 2007 Grand Challenge (27) found that, on average, interactive (i.e. semiautomated) segmentation methods were more accurate and reliable than fully automated methods. The larger standard deviation of automated methods was attributed to increased outlier errors (27).

A semiautomated liver segmentation method provides a suitable trade-off between full automation and manual segmentation. Such a hybrid method incorporates input from the user with a sophisticated understanding of liver imaging with automation of repetitive steps best performed by software. Human feedback can thus be provided at critical steps during segmentation to avoid error propagation.

A total interaction time of less than ten minutes for semiautomated liver segmentation would be ideal for practical purposes. Any task that requires a longer period away from clinical work would not be sustainable in the context of a busy radiology practice. Furthermore, sample size requirements to use parametric statistical tests and to obtain sufficient power in clinical studies typically require at least thirty patients. The manual segmentation of a liver can surpass 30 minutes per case. Thus, lengthy segmentation solutions may extend time required in validation of new segmentation methods and make them non-feasible. In our experience, a segmentation method that requires 5-10 minutes of total interaction time offers the right balance of user feedback and software automation.

A final clinical need we wished to address was the ability to segment both CT and MRI images using the same automated method. Though a vast amount of

literature exists regarding CT-based segmentation, published automated MR-based segmentation methods are limited as they are more difficult to develop and validate (90). Even more scarcely described in the clinical literature is a method compatible with both modalities. The feature-matching step **(Section 3.2.2)** which identifies the actual liver boundary for both CT and MRI images established the multi-modality versatility of our method. Section 5 describes a validation study for our segmentation method using CT and MRI images.

3.1.2 Quantitative Output

In order to generate uniform segmentation results which are comparable, segmentation rules should be instituted and clearly explained to image analysts. General consensus is to exclude major vessels which abut the liver periphery, such as the main portal vein and inferior vena cava. Other vessels which are completely surrounded by liver parenchyma (i.e. portal vein branches, hepatic veins and hepatic arteries) are generally included in the segmentation. Tumours and other pathologic structures peripherally located in the liver parenchyma are also usually included. As we will discuss in **Section 3.1.3**, at times it may be difficult to distinguish the boundaries between liver and surrounding structures due to similar tissue density, leading to segmentation error.

Once the liver has been completely segmented, the software should provide whole-liver volume with minimal input. In addition, the 3D segmented liver envelope should be exported and saved as a mesh that can be read by other software. Comparison of surface meshes provides an additional tool to compare segmentation accuracy both visually and with mathematical concepts.

Recently, segmentation evaluation frameworks have been criticized for using liver volume alone to evaluate the quality of segmentation results (27). To facilitate the comparison between segmentations and objectively assess technical improvements from different research groups, several performance measures have

been proposed by the liver segmentation community to highlight different aspects of segmentation agreement. These measures are summarized in **Appendix 1**.

3.1.3 Sources of Error

There are multiple potential sources of error which must be accounted for in the development of an automated liver segmentation method. In this context, error is defined as any element which impedes a segmentation method from producing exact results. This section describes potential sources of error related specifically to method initialization, choice of imaging modality, patient anatomy and evaluation frameworks. Possible solutions are also provided.

Error linked to method initialization

The susceptibility of a segmentation algorithm to certain sources of error is intrinsically linked to the type of algorithm used and its limits of use. Methods based on deformable models, such as the one described in this dissertation, usually contain three main steps: initialization of the model, identification of certain image characteristics and model deformation. The initialization step is classically the one most prone to error.

The deformable models approach requires the user to manually input an initial solution which can range from rough to precise. The more elaborate the initialization, the less the possibility of segmentation divergence and error. However, intricate initializations are lengthy to implement thus compromising efficiency.

The initialization of a surface model may be performed via a simple sphere within a region of interest or from numerous organ contours leading to a detailed model. When the morphological variability of an organ is low, such as for a bony structure, a low input initialization sequence may be sufficient. This would not be adequate for an organ such as the liver which demonstrates significant variability

amongst patients. In addition, it may be affected by numerous pathological conditions (i.e. cirrhosis, tumours) which cause its surface to be irregular.

On CT, liver density is usually relatively distinct from those of neighboring structures. For this reason, use of a classical technique such as region-growing may be considered for initialization. However, at times the liver parenchyma is indistinguishable or of similar density to adjacent organs. In these instances, region-growing may fail; either extending to adjacent organs or incompletely encompassing the liver parenchyma **(Figure 3.2)**. On MRI, lack of a clear density interface between the liver and adjacent structures also makes the use of region-growing problematic.

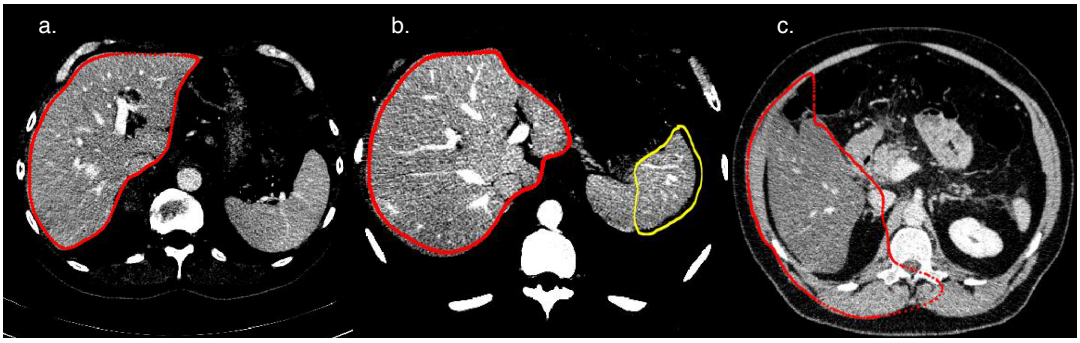

Figure 3.2: Initialization errors
Use of a region-growing technique to initialize CT-based automated segmentation may result in (a) adequate initialization, (b) extension to adjacent organs such as the spleen, (c) complete failure without adequate delineation of the liver. Figure used with permission from Gabriel Chartrand.

Our solution to the initialization problem involves manual delineation of the liver with six drawn contours. Two contours are placed per orthogonal plane in such a way to globally outline the liver contour while being sufficiently apart to capture specific hepatic features. In our experience, this number of contours is sufficient to generate a reliable initial shape, findings corroborated by Wimmer et al. (105).

From these initial sparse contours, it is possible to interpolate a 3D surface model using variational shape interpolation (106). This method is compatible with both CT and MRI images. Also, it allows the software to generate an initial shape quickly without the use of statistical shape models which have inherent limitations (described in **Section 2.5.3**). The initial shape is formed independent of image data and instead depends on user input which is assumed to more exact. As such, the initialization becomes robust and converges to the segmentation solution more rapidly.

Error linked to modality

The modality used for segmentation purposes dictates the quality of features which can be extracted from images. The more difficult it is to extract a coherent feature, the higher the risk of segmentation divergence. Image acquisition parameters and imaging artifacts directly influence segmentation results and represent important potential sources of error.

Liver CT and MRI acquisitions are typically performed with breath-hold to limit the effects of respiration on image quality. CT offers rapid image acquisition and thus excellent spatial resolution. MRI is known for longer image acquisition and requires a compromise between spatial resolution, signal-to-noise ratio and acquisition time. Slice thickness may be increased to obtain adequate z-axis coverage of the entire liver to accommodate the limited breath-hold capacity of some patients. This results in partial volume effects when voxels located at the interface of two structures with different signal characteristics must be averaged. In addition, large spaces between voxels need to interpolated making the volumetric calculations more prone to error.

Slice thickness has also been shown to impact automated liver volumetry results (79, 86, 107). Sahin et al. show that 4-5mm thick slices are most suitable to accurately estimate liver volume using MR imaging as compared to 2.5, 7.5 and 10mm thick slices (86). Reiner et al. suggest 6mm slice thickness on CT and 8mm

on MR provided a reasonable trade-off between volumetric precision and time efficiency (107). Our validation study in **Section 5** used average slice thicknesses of 2.5mm for CT and 5.5mm for MR. We believe these were optimal for the purposes of validation of our semiautomated method.

Use of volumetric imaging (3D MRI) allowed us to overcome errors associated with large slice thicknesses and gaps between slices. Rofsky et al. introduced 3D spoiled gradient echo sequences in 1999 (108). These MR pulse sequences allow near isotropic 3D imaging of the liver in one breath hold resulting in images with high spatial resolution. For our study we used a commercial sequence known as Liver Acquisition and Volume Acquisition (LAVA) sequence based on earlier 3D spoiled gradient-recalled echo sequences.

Both CT and MRI are prone to imaging artifacts, or elements within an image which do not represent normal patient anatomy and impact image interpretation. Artifacts can negatively influence automated segmentation methods causing segmentation error.

On CT, metallic objects such as surgical material can cause streak artifacts. These are typically more pronounced on MRI as metallic objects cause "blooming" when they locally influence the magnetic field. Blooming artifact can distort the image at the liver boundary causing segmentation error in these locations. Motion, pulsation and partial volume artifacts have also been shown to interfere with segmentation accuracy (79, 90). Artifacts and imaging pitfalls commonly seen on MRI are displayed in **Figure 3.3**.

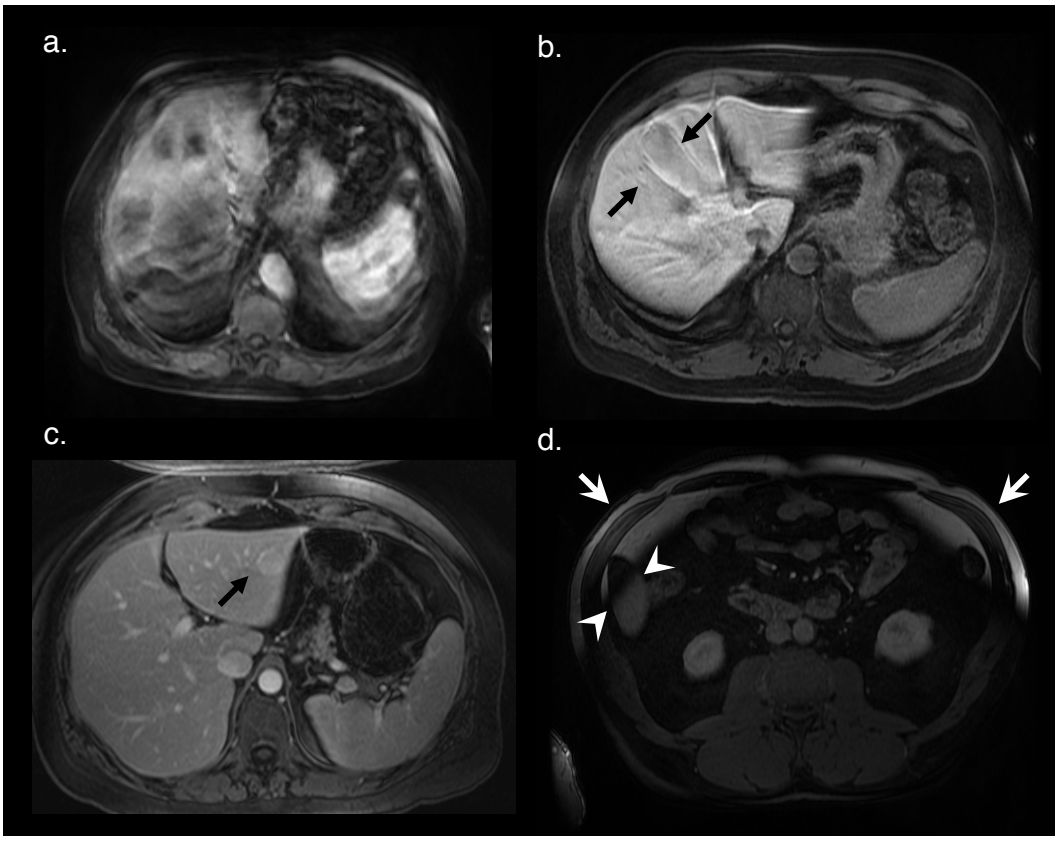

Figure 3.3:. Imaging pitfalls which may degrade liver segmentation on MRI. Axial T1-weighted fat-saturated images with contrast injection depict: (a) severe motion artifact, (b) partial volume averaging of the liver parenchyma with the gallbladder *(arrows)*, (c) ghost artifact with the aorta *(arrow)*, (d) inhomogeneous fat saturation *(white arrows)* and fat-water swap in the liver *(arrowheads)* (8).

The deformable models approach offers good resistance to noise and metallic artifacts on CT due to iterative and incremental rigid transformations. Artifacts on MRI which obscure the liver boundary are also overcome by conserving rigid boundaries. In addition, surfaces can be corrected directly by the user in areas degraded by artifact.

Error linked to anatomy

Certain liver anatomical characteristics inevitably cause segmentation difficulties. Segmentation error on CT images is often noted at the interface of liver parenchyma and the adjacent intercostal muscles, diaphragm, spleen, stomach and heart **(Figure 3.4)**. Review of the CT segmentation literature corroborates error at low contrast boundaries (27), at the liver hilum, adjacent to tumours, at hepatic fissures and near vascular insertions.

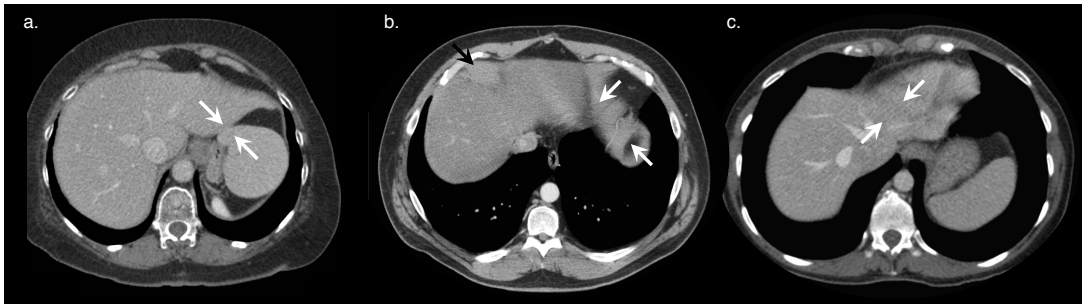

Figure 3.4: Imaging pitfalls which may limit liver segmentation on CT.
(a) Axial enhanced CT image of a 62-year-old woman shows indistinct liver-spleen boundaries (arrows). (b) Axial enhanced CT image of a 47-year-old man depicts segmentation challenges caused by ill-defined and non-continuous borders found near the liver dome (arrows). (c) Axial enhanced CT image of a 73-year-old man shows partial volume averaging between the left liver and the heart (arrows) (8).

Similar areas cause segmentation error on MRI. Under-segmentation on MRI occurs at low-contrast liver boundaries and areas of inhomogeneous density whereas over-segmentation usually occurs at organs abutting the liver (90).

Cirrhosis, steatosis, polycystic diseases, regions of ablation and malignancies (particularly tumours at the liver border) can cause highly irregular liver morphology. This can affect automated segmentation results, stifling the evolution of a deformable model to a complex form due to multi-lobulated contours.

Error associated with complex morphologies can be overcome by increased user interaction. Implementing more complex initialization strategies such as increasing the amount of detail per contour may improve segmentation accuracy in these difficult areas. Deformable models are not typically adapted to propagate into small regions such as hepatic fissures or vascular insertion sites. Error at these sites may be rectified by interactive correction tools, as are described in **Section 3.2.3**.

3.2 Segmentation Method

The proposed method was developed at the Laboratoire de recherche en imagerie et orthopédie (LIO, Montreal, Canada) with collaboration from the clinical and engineering teams. The code was implemented in C++ using VTK (Kitware Inc., 2014, Clifton Park, NY) as a rendering external library.

The semiautomated segmentation method consists of 3 main phases **(Figure 3.5)**. First, a 3D surface mesh is interpolated from a few user-generated contours of the liver. For each vertex of the generated mesh, matched features corresponding to the liver boundary are identified in the dataset. A Laplacian mesh optimization then deforms the mesh toward the matched features, while preserving surface smoothness. Finally, the user can manipulate and correct the final mesh.

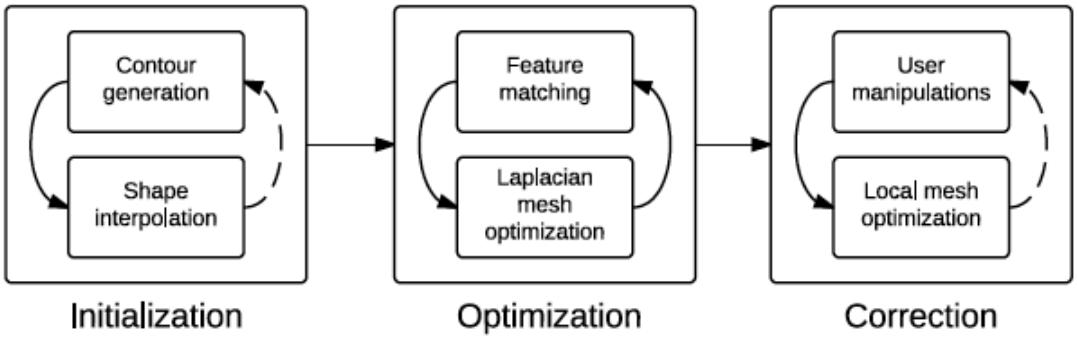

Figure 3.5: Proposed segmentation method.
The proposed segmentation method is composed of three phases: initialization, optimization and correction. Figure used with permission from Gabriel Chartrand and authors of (109).

3.2.1 Initialization Phase

The aim of the initialization phase is to provide a tool for the user to easily generate a reliable initial solution from as few interactions as possible. During this first step, the user is presented a graphical interface with traditional multi-planar views (i.e. axial, coronal, sagittal) and a 3D projective view where all drawn contours can be visualized. The user then clicks in the selected view to position nodes around the liver contour which are automatically connected by a cardinal spline **(Figure 3.6A)**.

Once the contour is closed, nodes can be moved, removed or added until the contour satisfactorily outlines the liver shape in the chosen view. Furthermore, the contours are automatically optimized using image warping and a minimal path algorithm, inspired from (110), to precisely delineate the liver boundary **(Figure 3.6B)**. This automatic contour optimization step differed slightly for CT and MRI images and is further described in Chartrand et al. (109). Positioning 2 contours per orthogonal plane generally provide enough constraints to produce an adequate initial solution. If needed, further contours can be added after the following interpolation step if the resulting shape is too far from the desired segmentation.

Variational shape interpolation

Variational shape interpolation (93) is a powerful shape interpolation method based on radial basis function (RBF) interpolation. This method is applied to generate a 3D surface mesh from these initial sparse contours.

As previously described by Heckel et al. (93), variational shape interpolation is a method which interpolates a 3D function of the form $f(x) = V$, where $x = \{x; y; z\}$, implicitly embedding the interpolated shape at its zero level-set. To obtain this function, a set of valued interpolation nodes c_i need to be defined on the surface of the target shape, as well as inside and outside the shape. The value $f(c_i) = 0$ is assigned to nodes lying on the desired surface, while interior and exterior nodes are assigned negative and positive values respectively, in such a way that the resulting iso-surface $f(x) = 0$ outlines the shape's boundary.

To better approximate the liver boundary orientation, we rely on the image information underlying the provided contours to define the oriented interpolation nodes. Since these were automatically optimized to fit onto the liver boundary, the underlying gradient information is assumed to globally represent the liver boundary. Every provided contour is regularly sampled at a given step. For each of these samples, the gradient orientation of the image data is estimated. For the gradient orientation to be robust to image noise and adequately represent the liver boundary normal direction, the image data is smoothed by a Gaussian kernel prior to sampling the gradient. The sampled gradient orientation is then inspected for inconsistencies, such as normal flipping, and further smoothed along the contour path to ensure smooth and continuous values **(Figure 3.6C)**.

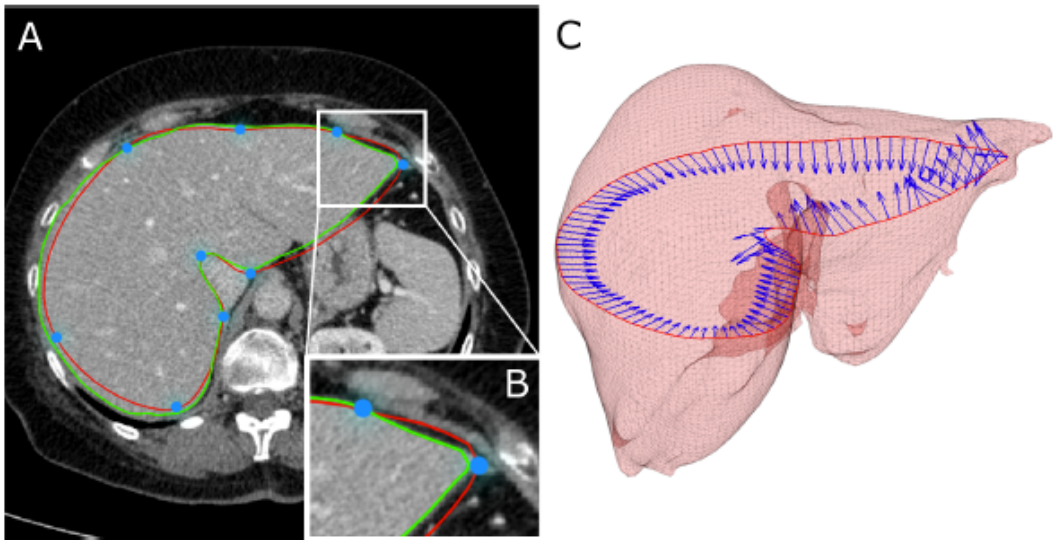

Figure 3.6: Initialization phase.
(A). The input spline in **red** generated with nodes in **blue** and optimized contour in **green**. (B) Detail of contour optimization. (C) Surface normals (blue arrows) of a single contour (red spline) estimated from image gradient information. Desired shape overlaid in pale red. Figure used with permission from Gabriel Chartrand and authors of (109).

Additionally, estimated 3D normals deviating from its associated projected contour normal for more than 45° are discarded. Finally, for each oriented sampled nodes, interpolation nodes are translated inward and outward along the final gradient orientation **(Figure 3.7A)**. Once the weights are computed, the resulting function implicitly defines a 3D manifold that intersects every previously defined contours. The final surface can then be extracted by evaluating the function over the domain of interest and then using an iso-surface algorithm to generate the surface mesh **(Figure 3.7C)**. Mathematical equations for variational shape interpolation are beyond the scope of this thesis, but are described in Chartrand et al. (109).

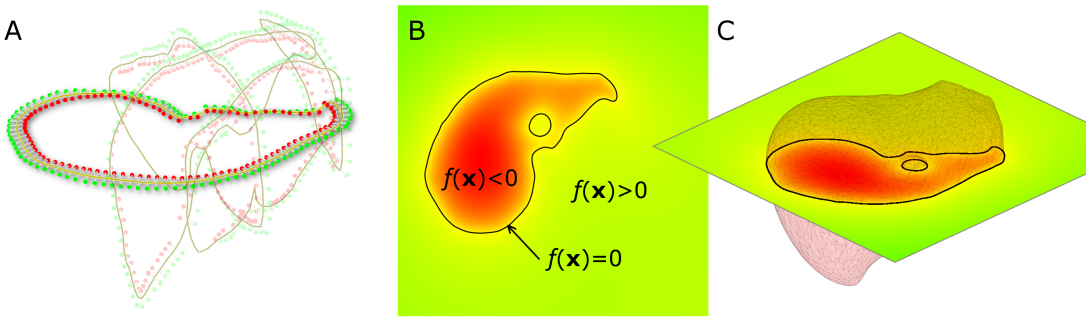

Figure 3.7: Variational shape interpolation.
(A) Variational shape interpolation rely on a set of sparse nodes with negative values inside the desired shape (red dots), positive values outside (green dots) and null values at the interface (yellow lines). (B) The resulting interpolation function is evaluated on a lattice from which a surface mesh is extracted using an iso-surface algorithm (C). Figure used with permission from Gabriel Chartrand and authors of (109).

3.2.2 Shape Deformation phase

Once the liver shape has been properly initialized, it is close to the desired segmentation but does not precisely overlap. This step is based on a Laplacian mesh optimization method (98), used as an iterative non-rigid registration approach to segmentation.

Feature matching

Feature-matching assigns each vertex of the initial 3D surface mesh with a corresponding target point representing the most probable location of the liver boundary. This target point is determined along intensity profiles as the point of maximal intensity difference between inward (liver) and outward (non-liver) intensities. This step differs for CT **(Figure 3.8A)** and MRI **(Figure 3.8B)** images, thus establishing the multi-modality versatility of the segmentation method.

For CT, since most of the liver parenchyma lies within a specific intensity range, the intensity profiles (gray signal in **Figure 3.8C**) are rescaled according to

a Gaussian transfer function (blue signal in **Figure 3.8C**) parameterized by the estimated mean intensity and standard deviation of the liver. After this operation the liver boundary, if apparent along the rescaled intensity profile, should present itself as a signal jump for normals pointing inward. To identify this feature, we rely on the sum of absolute differences similarity measure.

On MRI, rescaling the interpolated intensity profiles by a Gaussian transfer function would not be helpful due to intensity inhomogeneities across the liver parenchyma. However, since the liver is enhanced with a contrast agent in a late acquisition phase, its intensity is likely different from adjacent tissues though not entirely constant.

In this context, a metric inspired from Chan and Vese (111) was used where the difference between the mean intensity before and after the target feature of the profile is maximized. However, the maximum of the resulting signal, as seen on the blue signal in **Figure 3.8D**, is not always located at the liver boundary since the fat/muscle interface might display a strong intensity difference.

The obtained signal is therefore multiplied by the profile intensity (gray signal in **Figure 3.8D**) to favour bright features over darker ones. The signal is further multiplied by the profile's intensity derivative (green signal in **Figure 3.8D**) to favour edges as well. The target feature is then obtained from the maximum value and its location on the resulting signal (red signal in **Figure 3.8D**).

At each iteration, the shape progressively converges toward the desired segmentation and thus the length of the search space is linearly reduced to account for the mesh closing in on the patient's anatomy. Though the MRI feature matching method is prone to be noisy as the gradient is multiplied, our results generated few outliers and the large majority of our segmentations converged appropriately.

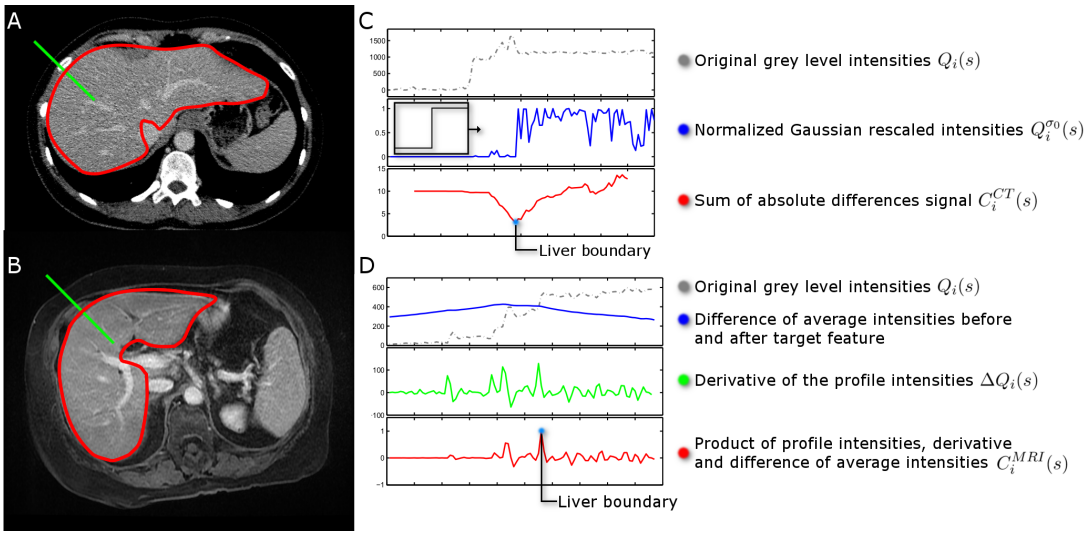

Figure 3.8: Feature matching strategy

The feature matching step identifies along intensity profile (A,B; green lines) the most probable location of the liver boundary. (C) For CT, a signal jump window is swept against the Gaussian rescaled profile to compute the sum of absolute differences similarity metric. (D) For MRI, the difference between the average intensity before and after every possible feature position along the profile is computed and weighted by the profile's intensity derivative. Figure used with permission from Gabriel Chartrand and authors of (109).

Laplacian mesh optimization

After the matching step, every vertex is assigned to a target with a certain confidence weight. While the target features might globally correspond to the liver boundary, they are most likely not assigned to their optimal location relative to one another. Laplacian mesh optimization (98) ensures that vertex relocation preserves a smooth local curvature while deforming the mesh toward their matched features. Mathematical equations for variational Laplacian mesh optimization are beyond the scope of this thesis, but are further described in Chartrand et al. (109).

3.2.3 Interactive Corrections Phase

At times, the initial 3D surface mesh might be too distant for the intensity profile to reach the liver boundary. Additionally, adjacent structures may display similar intensity to the liver parenchyma leading to target error. To recover cases where the segmentation might diverge from the patient's anatomy, two interactive corrections tools were implemented to impose additional constraints on the shape evolution. No regulation with respect to prior shape anatomy was used in this phase. The tools described below were intended for small corrections only.

The first correction tool **(Figure 3.9A)** allows the user to manipulate the surface mesh from the multi-planar view of the user interface. This tool allows the user to click on the surface mesh and manipulate it to the desired location. After releasing the mouse button, a locally constrained version of the Laplacian mesh optimization is launched.

The second correction tool **(Figure 3.9B)** allows the user to input constraint curves, the same way as in the initialization phase, to prevent the shape from evolving toward misleading features. If such constraints are defined prior to shape optimization, the curves are regularly sampled and each of the resulting constraint node is matched with the closest vertex on the surface mesh. This correction tool is particularly useful where large areas are missing clear boundary information or showing confusing features.

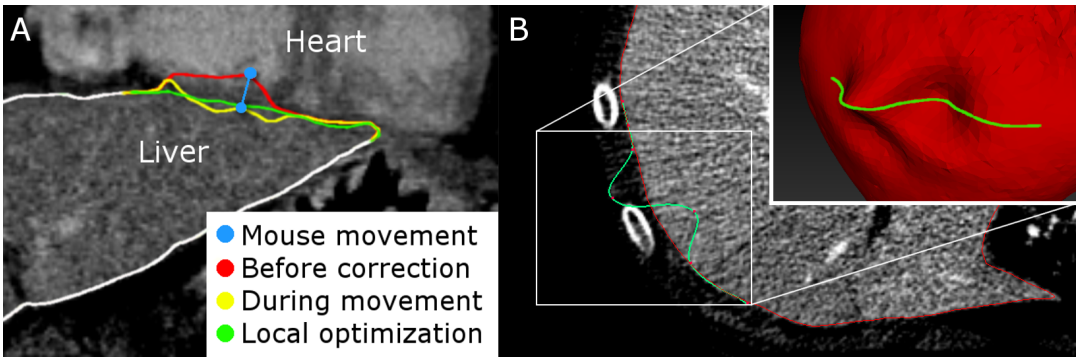

Figure 3.9: Interactive correction tools
(A) The first correction tool allows the user to manipulate the surface mesh from the MPR views. (B) The second tool allows the user to input strong positional constraint curves. Figure used with permission from Gabriel Chartrand and authors of (109).

3.3 Software Validation Strategy

As part of our research program, we devised a strategy to clinically validate our semiautomated liver segmentation method. The validation was performed *in vivo* with manual segmentation used as the reference standard. We were fortunate to have diverse patient databases representing a spectrum of liver disease available for our validation purposes.

As a first step, we introduced our novel semiautomated segmentation method as a proof of concept study. To study the method's multi-modality versatility, it was tested on diverse CT and MRI datasets. Semiautomated and manual segmentation were compared using segmentation performance measures which highlight various aspects of segmentation agreement. This study will be described in **Section 3.4**, summarizing the findings of a paper by Chartrand *et al*. (109).

As a second step, we evaluated the repeatability, agreement and efficiency of our method on a database of 41 subjects who underwent major hepatectomy between October 2006 and April 2009 and had a pre-operative contrast-enhanced CT. Segmentation quality was evaluated using segmentation performance

measures. This study will be described in **Section 4**, summarizing the findings of a paper by Gotra *et al*. (112).

As a third step, we compared the repeatability, agreement and efficiency of liver MRI- and CT-based semiautomated segmentation. The validation database consisted of 31 subjects with a spectrum of liver disease who required preoperative evaluation with both CT and MRI within two weeks between January 2010 and March 2013. Segmentation quality was once again evaluated using segmentation performance measures. This study will be described in **Section 5**, summarizing the findings of a paper by Gotra *et al.* (113).

3.4 Semiautomated Liver Segmentation on CT and MRI

In this section we summarize the findings of a paper by Chartrand et al. entitled "Liver Segmentation on CT and MR using Laplacian Mesh Optimization" (109).

As co-author for this manuscript, I was an active member of the research team responsible for creating the semiautomated liver segmentation method described. I provided clinical perspective and expertise during the developmental stages of the method. I participated in both the manual and semiautomated segmentation steps required for validation. Finally, I participated in the revision steps during manuscript drafting. This paper is currently being finalized prior to submission in a biomedical engineering journal not yet determined.

3.4.1 Introduction

Automated liver segmentation is a challenging task in the field of medical image processing. Usually performed on contrast-enhanced CT images, it provides physicians with 3D models and precise regions of interest for the evaluation of

numerous clinical parameters relevant in virtual surgery planning, radio-therapy planning and image-guided surgery.

The development and validation of automated liver segmentation methods represents a very active research area. The liver is an organ associated with various segmentation pitfalls. It can appear poorly contrasted on CT and MRI images **(Figure 3.10A)**, and is often in contact with adjacent organs which have the same image texture such as the spleen **(Figure 3.10B)**, the heart **(Figure 3.10C)**, and the stomach.

The shape of the liver varies considerably from one patient to another and its appearance is additionally variable depending on the medical or surgical history. The liver comprises intricate details such as vascular insertions and hepatic fissures which are difficult for automated algorithms to master.

Furthermore, imaging artifacts resulting from uneven contrast diffusion on CT or intensity inhomogeneities on MRI **(Figure 3.10D)** can impair automated segmentation processes. All these pitfalls combined makes it difficult to design an automated tool that is functional in every situation. Therefore, clinicians often fall back to manual segmentation though it cumbersome and not adapted to clinical reality.

We present in this work a semiautomated segmentation compatible with both CT and MRI images. The process is quickly initialized by the user drawing a few contours on multi planar views to globally outline the liver shape. A 3D surface model is then interpolated and automatically optimized to best fit image features. Two correction tools were also implemented to further correct the liver model until satisfaction. The proposed segmentation method, which was tested on CT and MRI datasets, was thoroughly analyzed and compared to contemporaneous methods.

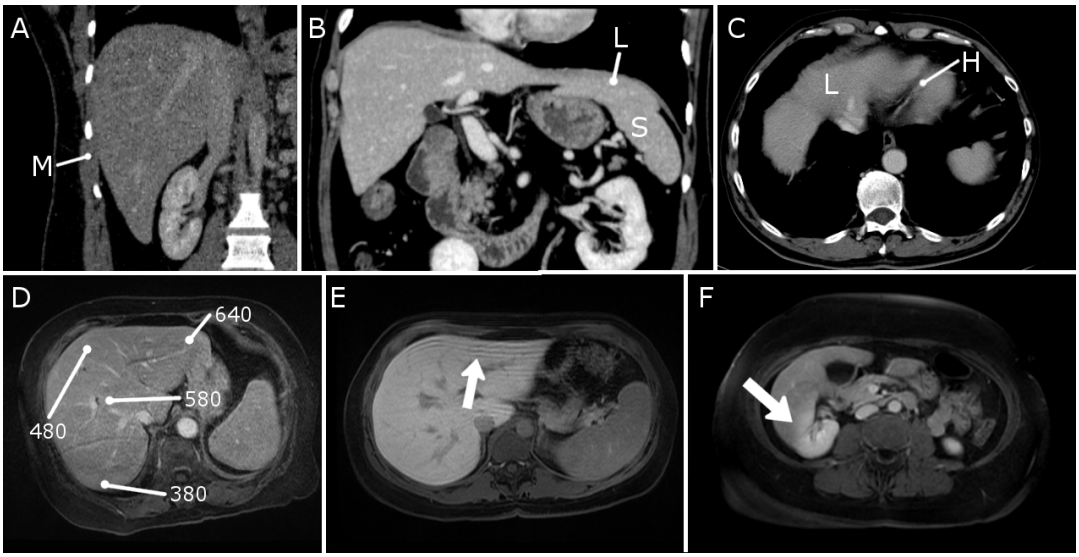

Figure 3.10: Common liver segmentation pitfalls on CT and MRI.
(A) Poor contrast with intercostal muscles (M). (B) Contact with the spleen (S). (C) Contact with the heart (H). (D) Variable signal intensity due to field inhomogeneity (measured intensity values in white). (E) Various acquisition artifacts such as truncation artifacts or (F) partial volume effect with adjacent organs (right kidney) due to large slice thickness. (L=Liver, M=Muscle, S=Spleen, H=Heart). Figure used with permission from Gabriel Chartrand and authors of (109).

3.4.2 Materials and Methods

The proposed semiautomated segmentation method being evaluated in this study was described in detail in **Section 3.2**.

Thirty CT examinations were obtained from the SLIVER07 repository, a common database of contrast-enhanced CT images available for training purposes (27). Following approval from the institutional review board, 20 MRI examinations from patients being referred for hepatic surgery with a variety of liver pathologies were also acquired. CT images in the portal venous phase and MRI 3D Liver Acquisition with Volume Acceleration (LAVA) sequences with contrast injection were selected for segmentation purposes.

Reference segmentations were generated by a radiology resident using slice-wise manual segmentation with the "snake" tool using SliceOMatic 4.3 Rev-11 software (TomoVision, Montreal, Canada). Semiautomated segmentations were successfully performed for all available CT and MRI datasets.

Several segmentation performance measures have been proposed by the liver segmentation community to highlight different aspects of segmentation agreement (definitions provided in **Appendix 1**). These were used to compared semiautomated and manual segmentation results.

3.4.3 Results

For the semiautomated method average initialization time was 115 seconds. Average optimization time was 60 seconds for CT images and 15 seconds for MRI images. Average interactive correction time was 180 seconds.

Average performance error measures for CT and MRI-based segmentation are provided in **Table III.II**. These are compared to results from other published methods. Most of the results reported on CT data were previously obtained from the SLIVER07 challenge test data. The MRI-based methods were tested on varying datasets

Table III.II: Comparison of segmentation performance measures

CT				**MRI**			
	VOE	*RVD*	*ASD*		*VOE*	*RVD*	*ASD*
Foruzan et al. (114)	8.3	1.8	1.4	Gloger et al. (115)	10.6	4.7	n/a
Peng et al. (116)	5.5	1	0.8	Huynh et al. (90)	12.0	3.6	n/a
Lopez-Mir et al. (94)	6.3	-2.4	0.8	Lopez-Mir et al. (104)	9.5	n/a	n/a
Maklad et al. (117)	5.8	-0.6	0.9	Siewert et al. (101)	n/a	4.2	n/a
Linguraru et al. (118)	8.0	2.2	1.4	**Proposed Method**	**7.6**	**1.6**	**1.5**
Beichel et al. (119)	5.2	1.0	0.8				
Freiman et al. (120)	8.6	2.8	1.5				
Kainmuller et al. (121)	7.0	-3.6	1.1				
Heiman et al. (122)	9.7	n/a	1.6				
Soler et al. (123)	n/a	n/a	2.0				
Aoyama et al. (124)	n/a	2.2	n/a				
Proposed Method	**5.2**	**1.1**	**1.0**				

VOE: Volumetric Overlap Error (%), 0% for a perfect overlap between segmentations.

RVD: Relative Volume Difference (mm), 0mm implies that the segmentation volumes are identical.

ASD: Average Symmetric Surface Distance (mm), 0mm implies perfect segmentation.

Examples of raw segmentation results following initialization and optimization phases with corresponding manual segmentations are provided in **Figure 3.11**.

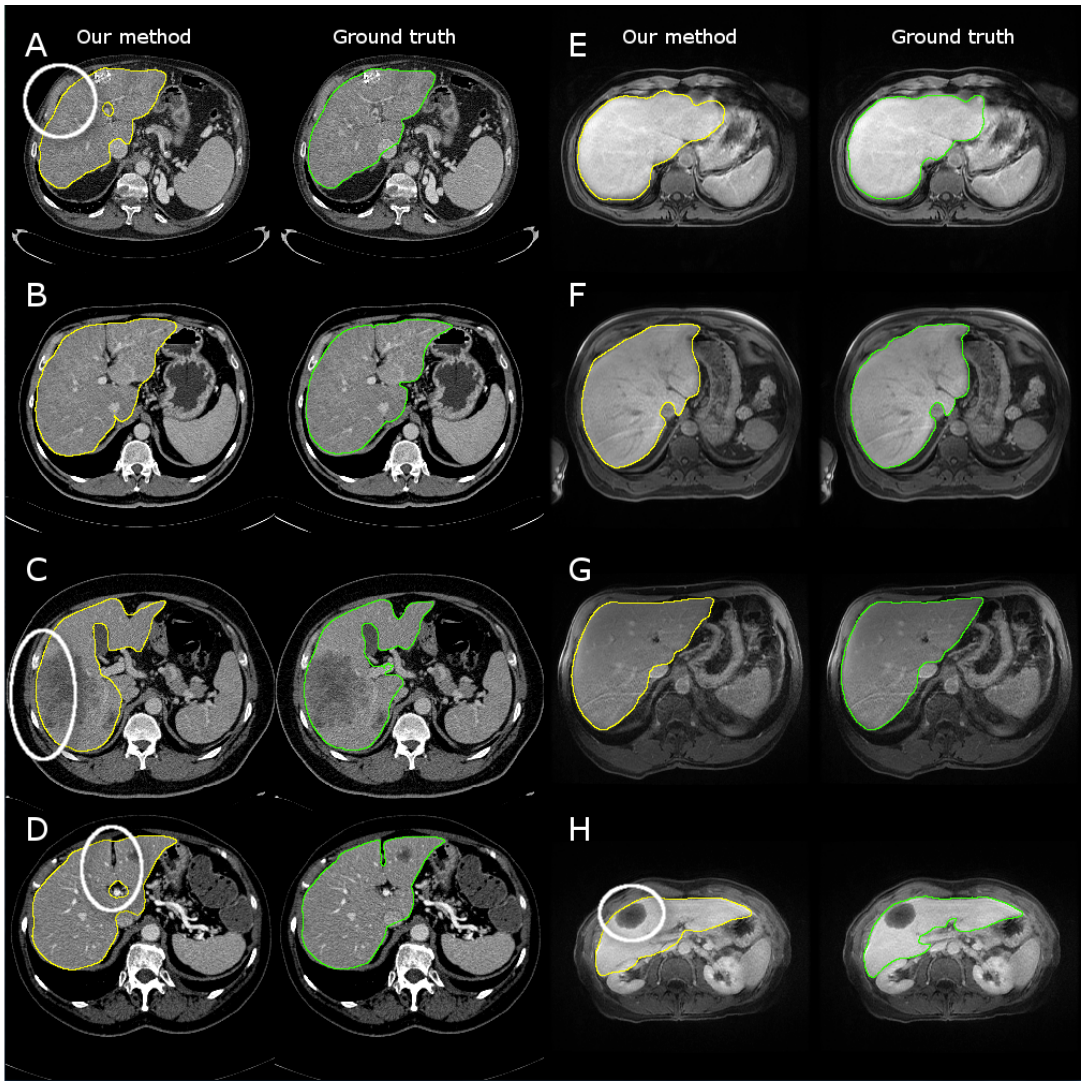

Figure 3.11: Examples of segmentation results.

Examples of segmentation results before manual corrections were made, with the corresponding ground truth segmentations on CT (A-D) and MRI (E-H). (A) The circle represents an area of the liver in contact with intercostal muscles which was successfully outlined. (B),(E), (F) and (G) display typically obtained results. (C) and (H) display peripheral tumours adequately included in the segmentation. (D) illustrates the difficulty of modeling thin and elongated features such as hepatic fissures. Figure used with permission from Gabriel Chartrand and authors of (109).

3.4.4 Discussion

Overall performances were inferior on MR images, which was to be expected due to signal heterogeneity, parallel imaging artifacts, susceptibility artifacts due to metal and air/tissue interfaces and partial volume effects due to slice thickness. These factors combined inevitably make manual and automated segmentation less accurate and prone to differ.

The segmentations on CT obtained using the SLIVER07 test data were submitted to the challenge organizers, which placed this method in 9th position amongst 80 contestant as of December 2014.

By visually reviewing the optimization results, we noted that the main areas of discrepancy were recurrent. The inclusion or the exclusion of the IVC often varies from one user to another. In the SLIVER07 database, the IVC was modeled in a way to preserve the continuity of the liver parenchyma's surface **(Figure 3.11A)**, whereas the proposed method lent itself better to a complete inclusion or exclusion. Similarly, the portal vein was often modeled differently than the reference with the proposed method **(Figure 3.11C)**.

Hepatic fissures led to important surface errors **(Figure 3.11D)**. Unless they were modeled initially, the rigidity of the surface mesh prevented it from propagating into these thin and elongated features. Poor contrast with intercostal muscle was well managed in general **(Figure 3.11A)** but failed completely in two particular cases which we considered outliers in this study.

The MRI feature matching strategy is implicitly less specific than the one applied for CT due to the inherent difficulty to model the liver appearance. Consequently, the convergence was less pronounced and the surface mesh was more easily attracted to erroneous features such as the anterior cortex of the right kidney **(Figure 3.11H)**.

Despite the minor errors reported, the proposed method overall performed well and achieved good results against common segmentation pitfalls such as

peripheral tumours **(Figure 3.11A)** and intensity inhomogeneities **(Figure 3.11G-H)** on MRI.

Beside satisfying the initial design objective of being free from training data and working on both CT and MRI, the method remained robust to most of the cases and provides intuitive and efficient correction tools to manipulate the segmentation until satisfaction. Moreover, the average segmentation time was reasonably under 5 minutes with an unoptimized implementation and free parameters being easily and instinctively set.

3.4.5 Limitations

We note however that our study involves some limitations. First of all, manual segmentation, commonly accepted as a gold standard surrogate in segmentation literature, inevitably lead to intra- and inter- reader variability on repeated segmentations, as discussed in a recent study (125). This is especially true for MRI liver segmentation, where partial volume effect are substantial, leading to some axial images being hard to interpret.

Furthermore, in some cases, a clear consensus regarding vessel and hepatic fissure exclusion can hardly be established. These areas often impact importantly surface distance performance metrics, even though they are less relevant toward the clinical outcome.

Additionally, since manual segmentation was supported by a 2D snake tool which evolves input contours on highest gradient, the segmentation was slightly overestimated due to some partial volume effect.

Finally, in the current implementation, the graphical user interface did not permit contour initialization in arbitrary slice orientations. This limited the possibility for the user to model in the initial surface mesh some structures such as hepatic fissures which are seldom aligned with the orthogonal viewing planes.

3.4.6 Conclusion

We present in this paper a semiautomated segmentation method that can be used for both CT and MRI liver segmentation. A primary aim was to overcome the need for training data while remaining robust and efficient on a wide range of pathological livers. Correction tools were implemented to provide the user the means to improve the segmentation until satisfaction.

Obtained results show that the Laplacian mesh optimization framework can achieve excellent segmentation in a short time with limited interaction. Adaptation to other MRI sequences as well as vascular subsegmentation will be addressed in the future.

4 Validation of a Semiautomated Liver Segmentation Method Using CT for Accurate Volumetry

As first author of this manuscript, I was involved in all aspects from study design to manuscript drafting. I was an active member of the team responsible for developing the liver segmentation method by providing clinical perspective. I participated in the ethics submission, raw data collection, manual and semiautomated segmentation steps for validation and statistical analysis. I led the manuscript drafting and revision process under the supervision of Dr. An Tang. This paper was accepted for publication in Academic Radiology in September 2015. For full text, please see **Appendix 2**.

4.1 Abstract

Rationale and objectives: To compare the repeatability and agreement of a semiautomated liver segmentation method with manual segmentation for assessment of total liver volume on CT (computed tomography).

Materials and Methods: This retrospective, institutional review board-approved study was conducted in 41 subjects who underwent liver CT for preoperative planning. The major pathologies encountered were colorectal cancer metastases, benign liver lesions and hepatocellular carcinoma. This semiautomated segmentation method is based on variational interpolation and 3D minimal path-surface segmentation. Total and subsegmental liver volumes were segmented from contrast-enhanced CT images in venous phase. Two image analysts independently performed semiautomated segmentations and 2 other image analysts performed manual segmentations. Repeatability and agreement of both methods were

evaluated with intra-class correlation coefficients (ICC) and Bland-Altman analysis. Interaction time was recorded for both methods.

Results: Bland-Altman analysis revealed an intra-reader agreement of -1 ± 27 mL; (mean ± 1.96 standard deviation) with ICC of 0.999 ($p < 0.001$) for manual segmentation and 12 ± 97 mL with ICC of 0.991 ($p < 0.001$) for semiautomated segmentation. Bland-Altman analysis revealed an inter-reader agreement of -4 ± 22 mL with ICC of 0.999 ($p < 0.001$) for manual segmentation and 5 ± 98 mL with ICC of 0.991 ($p < 0.001$) for semiautomated segmentation. Inter-method agreement was found to be 3 ± 120mL with ICC of 0.988 ($p < 0.001$). Mean interaction time was 34.3 ± 16.7 minutes for the manual method and 8.0 ± 1.2 minutes for the semiautomated method and ($p < 0.001$).

Conclusion: A semiautomated segmentation method can substantially shorten interaction time while preserving a high repeatability and agreement with manual segmentation.

4.2 Introduction

Assessment of liver volume is a mandatory step prior to extended hepatectomy for determining the anticipated future liver remnant and prior to living donor liver transplantation for selection of appropriate candidates (13, 72, 126). Liver volumetry requires a multiplanar imaging modality. CT is currently the preferred imaging modality for surgical planning due to its superior spatial resolution and short acquisition time (14, 22, 23). Use of CT in pre-surgical imaging allows for concomitant assessment of vascular anatomy and quality of liver parenchyma and allows determination of total and lobar volume (16).

The reference standard method to estimate liver volume involves manually delineating the liver outline, a process called “segmentation”, on consecutive CT images. This method is cumbersome, time-consuming and impractical for widespread clinical use (17, 127, 128). Formula-based liver volume estimation

using patient height and weight has also been proposed (129). However, this approach is based on a linear regression equation and is not specific to patient anatomy (130).

Automated segmentation algorithms provide several advantages such as shorter processing time, greater agreement and repeatability (18, 23, 79, 131). Although numerous studies have proposed semi- or fully-automated liver segmentation methods from CT datasets, these methods have not necessarily been translated to clinical use (27). Reasons limiting the performance of segmentation algorithms have included: small sample sizes, data sets not reflective of clinical problems and poorly defined performance metrics (27, 28). Recently, segmentation evaluation frameworks have been criticized for using liver volume alone to evaluate the quality of segmentation results (27). To facilitate the comparison between segmentation methods and objectively assess technical improvements from different research groups, several error measures have been proposed by the liver segmentation community to highlight different aspects of segmentation agreement: volumetric overlap error, average symmetric surface distance, root mean square symmetric surface distance and maximum symmetric surface distance (27).

4.2.1 Hypothesis

In this article, we introduce a novel semiautomated liver segmentation method for CT based on variational interpolation and minimal path surface segmentation. We hypothesized that this method would improve repeatability and agreement with manual segmentation while providing faster (i.e. more efficient) segmentation time. Our method is an improvement to previously published methods as no statistical shape model was imposed, which permits more segmentation flexibility for pathological or livers with unusual shape. This method is compatible with both CT and MR datasets, which has not been previously described to our knowledge.

Finally, the method is equipped with mesh-based correction tools which allow the user to achieve greater precision during interactive segmentation.

4.2.2 Aim

The primary aim of this study was to compare the repeatability, agreement and efficiency of a semiautomated liver segmentation method by using manual segmentation as the reference standard. A secondary aim was to evaluate the quality of segmentation using error metrics based on volume overlap and surface distances. Subsegmental volumetry was also performed based on vascular landmarks and classic anatomic principles defined by Couinaud (6).

4.3 Materials and Methods

4.3.1 Study Design

Our institutional review board approved this retrospective, cross-sectional study. Requirements for informed consent were waived.

4.3.2 Study Subjects

Our validation database consisted of 41 subjects (22 men, 19 women; mean age, 55 years) who underwent hepatectomy between October 2006 and April 2009 at our institution. Patients were included if they had primary or metastatic liver tumours and underwent major hepatectomy (≥ 3 Couinaud segments) during the study period. Each patient had a pre-operative CT scan within three months of surgery. Hepato-biliary surgeons at our hospital independently determined indications for pre-operative imaging according to clinical standard of care without influence for study inclusion. Study subjects' demographic and clinical information are summarized in **Table IV.I**.

Table IV.I: Subject demographics

Characteristic	Data
Total subjects, N (%)	41 (100)
Sex	
Male (%)	22 (54)
Female (%)	19 (46)
Age (y)	
Mean ± SD	55 ± 13
Body mass index in adults (kg/m^2)	
Mean ± standard deviation	26 ± 5
Pathologies	
Colorectal metastases	27 (66)
Hepatocellular carcinoma	4 (10)
Benign liver lesions	5 (12)
Biliary trauma	1/41 (2)
Cholangiocarcinoma	2 (5)
Cystadenocarcinoma	1 (2)
Cholangitis	1 (2)

4.3.3 CT Imaging Technique

CT was performed using two MDCT scanners under standard abdominal imaging protocols. Twenty-five study patients were scanned with a 16-detector scanner (Lightspeed 16, GE Medical Systems, Waukesha, WI) and 16 patients were scanned with a 64-detector scanner (Brilliance 64, Philips Medical Systems, Cleveland, OH).

The parameters for the 16-detector scanner were: rotation time, 0.8 seconds; detector collimation, 16 x 1.25 mm; helical pitch, 1.375; tube voltage, 120-140 kV; X-ray tube current: 75-440 mA; tube current–time product, 250 mAs. The

parameters for the 64-detector scanner were: rotation time, 0.75 seconds; detector collimation, 64 x 0.625 mm; helical pitch, 0.891; tube voltage, 120 kV; X-ray tube current: 151-499 mA; tube current–time product, varied based on noise index.

Image reconstruction was in a 282-500 mm display field of view, depending on the patient's physique. Reconstruction section thickness was 2.5 mm. Reconstructed CT slices had a matrix size of 512 x 512 pixels with pixel spacing ranging from 0.55-0.98 mm. Prior to all examinations, a weight-adjusted dose of a non-ionic, low osmolar, iodinated contrast agent (95-200 ml Isovue; Bracco Diagnostic Inc., Princeton, NJ) was administered intravenously with a 20-gauge needle at a rate of 4 ml/second. All CT protocols included an arterial phase and portal venous phase with delays of 40 seconds and 60 seconds respectively.

4.3.4 Study Workflow

The portal venous phase from the 41 subjects was used for segmentation as it provides homogeneous enhancement of the liver parenchyma and maximizes contrast between liver and non-liver structures. Liver segmentation (manual and semiautomated) was performed independently by four image analysts (one radiology resident, two medical students and one biomedical engineering PhD candidate) participating in research within the department of Radiology. Prior to this study, the image analysts received 10 hours of training in liver anatomy and software segmentation. Furthermore, the manual segmentation results used as the reference standard were supervised by an abdominal radiologist (7 years of experience).

Two image analysts performed manual segmentation while the other two undertook semiautomated segmentation. This ensured adequate estimation of agreement and intra- and inter-observer repeatability. Image analysts performed repeat segmentations in a random order one week later to prevent recall bias. Image analysts were blinded to the results of their first segmentation and to the

results of the other readers. Interaction time was recorded for both segmentation methods.

4.3.5 Manual Segmentation

Axial portal venous phase CT images for each patient were saved as DICOM files and uploaded onto an imaging post-processing display software (SliceOmatic 4.3 Rev-11, TomoVision, Montreal, Canada). For a given axial slice, two image analysts manually outlined the liver using a cursor to contour the liver. These curves then automatically deformed to precisely delineate the liver. This process generated "active contours" which are virtual curves that can be projected within images to delineate the liver boundary based on an energy equation (see **Section 2.5.1**) (91). Each axial slice required further manual deformation of the active contours to completely outline the liver. Large vessels abutting the liver periphery such as the main portal vein and inferior vena cava were excluded, but not vessels surrounded by liver parenchyma. The number of pixels within each contour provided the liver area on a slice-by-slice basis. This cross-sectional area was multiplied by the slice thickness and the summation of each section volume provided the total liver volume for each patient. Volumes and masks obtained from manual segmentation were used as the reference standard.

4.3.6 Semiautomated Segmentation and Subsegmentation

Our semiautomated segmentation method was developed at the Imaging and Orthopaedics Research Laboratory (LIO, Montreal, QC) with collaboration from the clinical and engineering teams. The method was developed and tested using MATLAB (2012a, The MathWorks, Inc., Natick, MA, USA) computational software. Axial portal venous phase CT images for each patient were saved as DICOM files and uploaded to the segmentation program. An overview of the user

and computer tasks involved in semiautomated liver segmentation is provided in **Figure 4.1**.

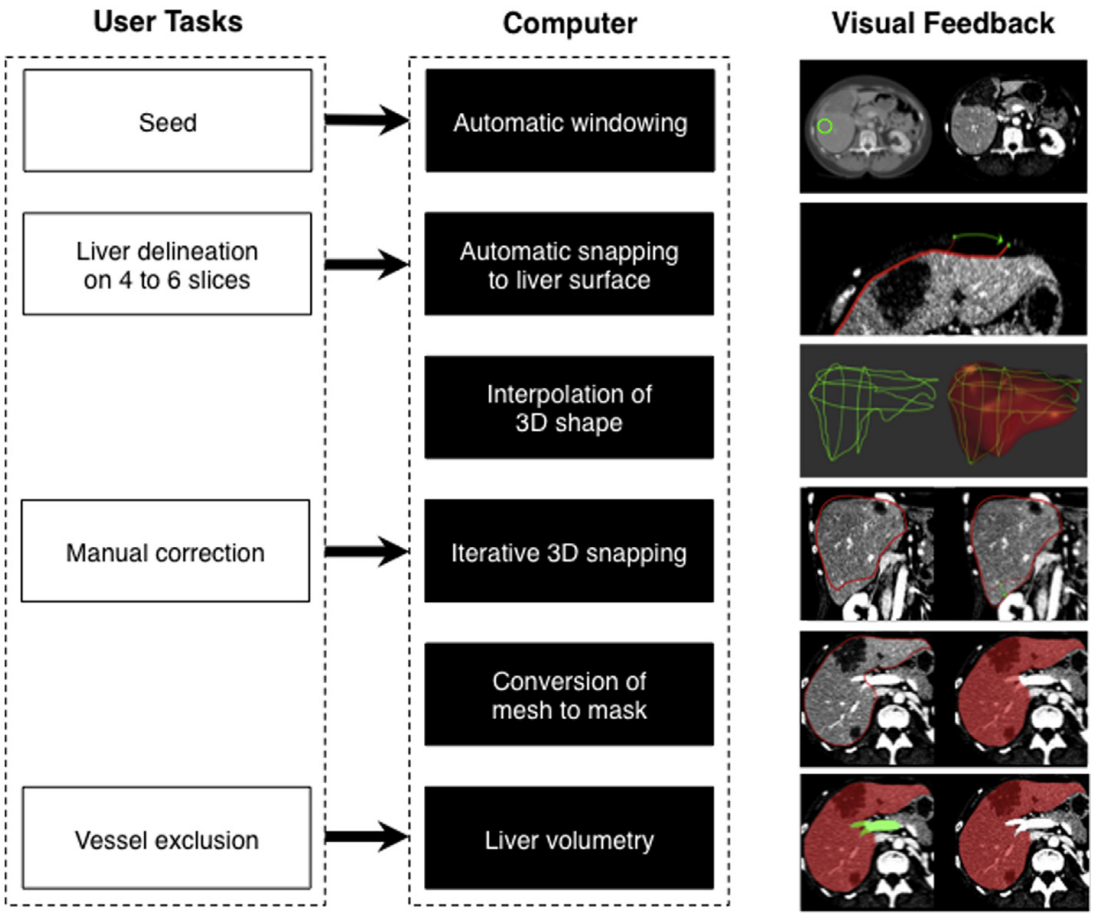

Figure 4.1: Overview of steps in CT-based semiautomated liver segmentation. The user initiates segmentation by roughly delineating the liver contour on 4-6 slices. The software then uses variational interpolation to generate an initial 3D shape. This 3D shape is deformed manually then automatically by minimal path surface segmentation. Vessels are excluded using a locally seeded region growing technique. The software then calculates liver volume for each slice.

Initially, a seed is positioned within the liver to define a volumetric spherical region of interest used to estimate the mean intensity and standard deviation. These values are used to automatically adjust the displayed contrast level and

windowing to enhance the liver boundary against adjacent tissue. A liver probability density map is then generated by applying a Gaussian transfer function.

In order to generate an initial shape without any prior knowledge, the liver is manually delineated on 1-2 slices for each orthogonal plane to globally outline the liver shape. This delineation process is assisted by a snapping algorithm based on image-warping and minimal path segmentation (132). As a result, the drawn contours dynamically snap onto the liver surface.

Variational interpolation is applied to these sparse contours to interpolate a smooth surface mesh composed of vertices and triangular faces intersecting the contours initially delineated (93, 106). In order to simplify the segmentation problem to a narrow band along the prior shape's surface, the mesh is further converted to a quadrangular mesh through surface parameterization (133). This allows the unfolding of the prior shape and the narrow band subspace which simplifies further segmentation.

The parameterized surface is then subject to two concurrent segmentations operations. First, the user can iteratively deform the mesh in 3D by adjusting the contours to align with actual liver anatomy. Second, the user can prompt an automated minimal surface segmentation technique to precisely delineate the liver boundary, a 3D extension of a method described by Chav et al. (132) .

The final segmented mesh is converted to a volumetric mask to exclude vessel insertion points and hepatic fissures with a local region growing tool. The cross-sectional area of each mask was multiplied by the slice thickness and the summation of each section volume provided the total liver volume for each patient.

For sub-segmentation, three vertical planes were defined by drawing lines through the left, middle, and right hepatic veins and their insertion at the inferior vena cava (IVC). The portal vein bifurcation established a horizontal plane to divide segments II/III, IVa/IVb, V/VIII and VI/VII. A polygon was then drawn to encapsulate liver tissue between the posterior aspect of the portal bifurcation and

the IVC. This polygon was propagated (using an automated tool) to other slices to define the caudate lobe **(Figure 4.2)**. Whole and segmental liver volumes are reported in **Table IV.II**.

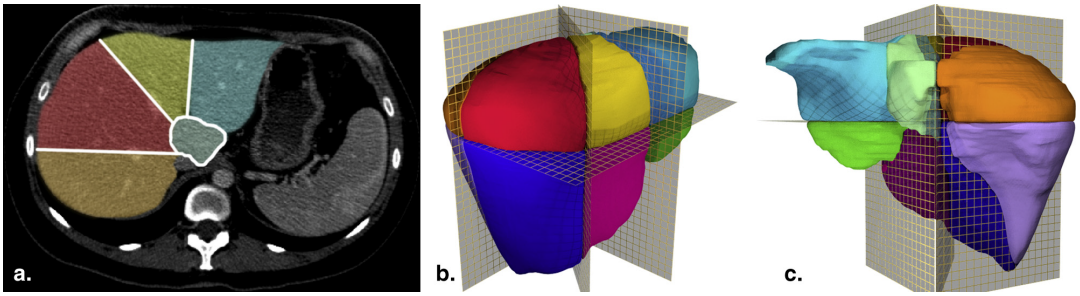

Figure 4.2: Liver Subsegmentation.

(A) Axial CT slice demonstrating caudate lobe, and segments II, IVa, VII and VIII. Three vertical planes are defined by drawing lines through the left, middle, and right hepatic veins and their insertion at the IVC. A polygonal shape is propagated to define caudate lobe. **(B)** and **(C)** Oblique anterior-posterior and posterior-anterior 3D renderings defining the liver subsegments.

Table IV.II: Whole and segmental liver volumes by reader.

Readers	**1 (Manual)**	**3 (Semiautomated)**	***p*-value[a]**
Whole liver volume (mL)[b]	1689 ± 478	1688 ± 497	0.92
Readers	**1**	**2**	***p*-value[a]**
Segmental volume (mL)[b]			
I	41 ± 16	53 ± 37	0.01
II	204 ± 110	186 ± 77	0.31
III	97 ± 66	74 ± 57	0.05
IVa	186 ± 77	205 ± 83	0.18
IVb	84 ± 54	101 ± 87	0.10
V	292 ± 99	278 ± 116	0.26
VI	221 ± 110	202 ± 114	0.14
VII	278 ± 106	292 ± 132	0.24
VIII	292 ± 103	306 ± 121	0.23

[a] Probability associated with a Student's paired t-test with a two-tailed distribution.

[b] Results reported as mean ± standard deviation.

4.3.7 Statistical Analysis

Statistical analyses was performed with SPSS software for Windows, version 21.0 (Chicago, IL). Whole and segmental liver volumes were compared using paired T-tests. Intra-class correlation coefficients (ICC) were used to determine intra-reader, inter-reader and inter-method variability of hepatic volume.

Bland-Altman analyses were used to determine intra-reader, inter-reader and inter-method agreement. The agreement for liver volume between readers and segmentation sessions was reported as bias ± 1.96 SD of the differences, followed

by the 95% limits of agreement interval (134). P-values were calculated for Bland-Altman analyses to evaluate for systematic bias different from 0.

A sub-group analysis was performed in patients with hepatocellular carcinoma (HCC). This analysis was done to determine whether the presence of underlying fibrosis or cirrhosis, which are risk factors for HCC development, affected the results of semiautomated liver volumetry.

The differences between semiautomated and manually segmented surface meshes were analyzed with 4 additional error measures: volumetric overlap error, average symmetric surface distance, root mean square (RMS) symmetric surface distance and maximum symmetric surface distance (27). The formulas to calculate these segmentation error measures are reported in **Appendix 1**.

In addition, paired T-tests were used to compare the total interaction time for semiautomated segmentation with manual segmentation time.

4.4 Results

4.4.1 Volumes

The mean semiautomated whole liver volume was 1688 ± 497 mL, whereas the reference standard volume was 1689 ± 478 mL (P = .92). Mean segmental volumes are demonstrated in Table 4.2. The only statistically significant difference when comparing segmental volumetry was for the caudate lobe (P = .01).

4.4.2 Variability

Overall 8 measurements of ICC representing intra-reader, inter-reader and inter-method variability of hepatic volume measurements were calculated, these are summarized in **Table IV.III**. Correlation was high with an agreement between the eight ICC measures of 0.995 (95% confidence interval [CI]: 0.992-0.997).

Correlation between semiautomated and manual volumetry was established with inter-method ICC values ≥ 0.988 (P< .001). Correlation for segmental volumetry readings varied greatly with values ranging from 0.331 (segment III) to 0.831 (segment VII).

Table IV.III: Intra-reader repeatability, inter-reader and inter-method agreement

Comparison	Readers	ICC[a]	Bland-Altman (mL)[b]
Repeatability on whole liver volume			
-Intra-reader manual	1 vs 1	0.999	-1 ± 27 (-28, 26)
	2 vs 2	1.000	-6 ± 11 (-17, 6)
-Intra-reader semiautomated	3 vs 3	0.995	-3 ± 67 (-70, 64)
	4 vs 4	0.991	12 ± 97 (-85, 109)
Agreement on whole liver volume			
-Inter-reader	1 vs 2	0.999	-4 ± 22 (-27, 18)
	3 vs 4	0.991	5 ± 98 (-93, 103)
-Inter-method[c]	1 vs 3	0.992	-2 ± 93 (-95, 91)
	1 vs 4	0.988	3 ± 120 (-117, 124)
Agreement on segmental volumes			
-Inter-reader			
Segment I	1 vs 2	0.585	12 ± 59 (-47, 71)
Segment II	1 vs 2	0.399	-17 ± 207 (-224, 190)
Segment III	1 vs 2	0.331	-23 ± 139 (-162, 116)
Segment IVa	1 vs 2	0.458	18 ± 164 (146, 182)
Segment IVb	1 vs 2	0.713	16 ± 121 (-105, 181)
Segment V	1 vs 2	0.758	-14 ± 150 (-164, 136)
Segment VI	1 vs 2	0.728	-20 ± 162 (-182, 142)
Segment VII	1 vs 2	0.831	14 ± 144 (-130, 158)
Segment VIII	1 vs 2	0.812	14 ± 139 (-125, 153)

[a] ICC = Intra-class correlation coefficient.

[b] Bland-Altman = Results reported as bias ± repeatability coefficient (1.96 SD); (95% limits of agreement interval), rounded to whole numbers.

[c] Inter-method agreement is reported 1 vs 3 and 1 vs 4, which represent the worst-case scenarios when comparing manual and semiautomated volumetry.

4.4.3 Repeatability

Bland-Altman analysis showed excellent repeatability for both manual and semiautomated CT-based volumetry **(Table IV.III)**. Intra-reader agreement for the manual method was -6 ± 11 mL with limits of agreement of -17 and 6 mL (P= .426). The semiautomated method displayed higher bias: 12 ± 97 mL and wider limits of agreement: -85 and 109 mL in the repeatability calculations (P= .291).

4.4.4 Agreement

Bland-Altman analysis showed good agreement between readers for each method and between methods **(Table IV.III)**. Inter-reader agreement for the manual method had a bias of -4 ± 22 mL with limits of agreement of -27 and 18 mL (P= .009). Inter-reader agreement for the semiautomated method had a bias of 5 ± 98 mL and limits of agreement of -93 and 103 mL (P= .293). Inter-reader agreement for segmental volumes demonstrated generally large limits of agreement; ranging from -47 and 71 (segment I) to -224 and 190 (segment II). These limits of agreement were wider than those for whole liver volumetry.

The agreement between manual and semiautomated volumetry methods was 3 ± 120 mL with limits of agreement of -117 and 124 mL (P= .434), represented in **Figure 4.3**. Examples of concordant and discordant cases between readers are demonstrated in **Figures 4.4** and **4.5** respectively.

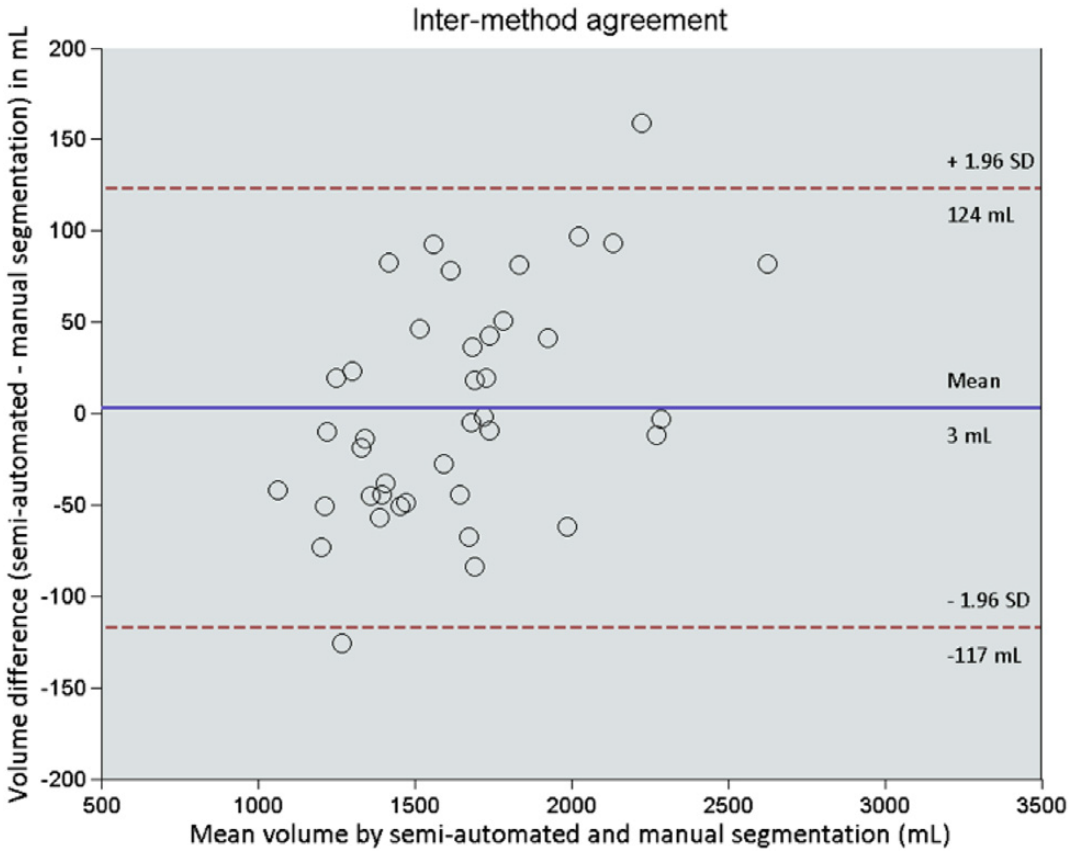

Figure 4.3: Inter-method agreement.

Bland–Altman plot of the volume difference between semiautomated and manual segmentation of computed tomography images and the mean volume (reader 1 vs. reader 4). Mean difference was demonstrated with solid line and 95% limits of agreement with dashed lines.

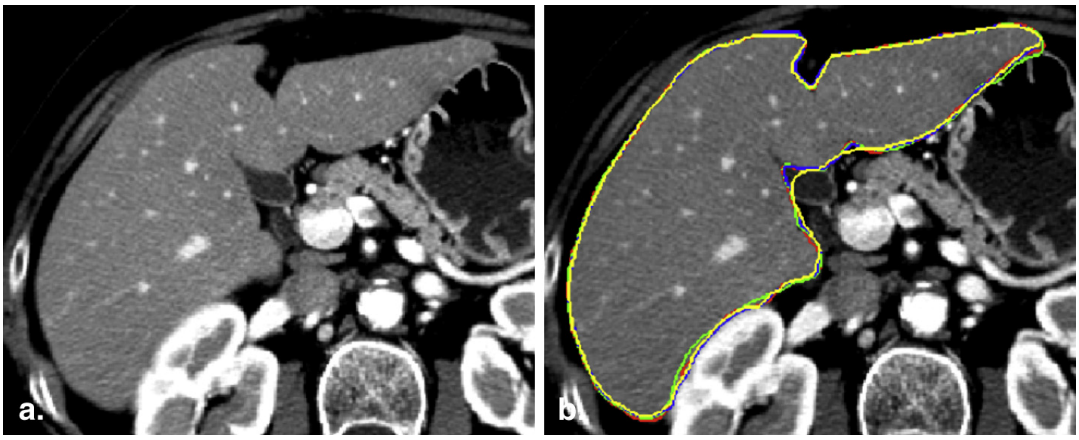

Figure 4.4: Concordant liver segmentation.

67-year-old woman with colorectal metastases. **A** and **B**, Original **(A)** and annotated **(B)** Axial CT slice demonstrating concordance between four readers using manual and semiautomated liver segmentation methods. Reader 1 manual = red tracing, reader 2 manual = green tracing, reader 3 semiautomated = blue tracing, reader 4 semiautomated = yellow tracing.

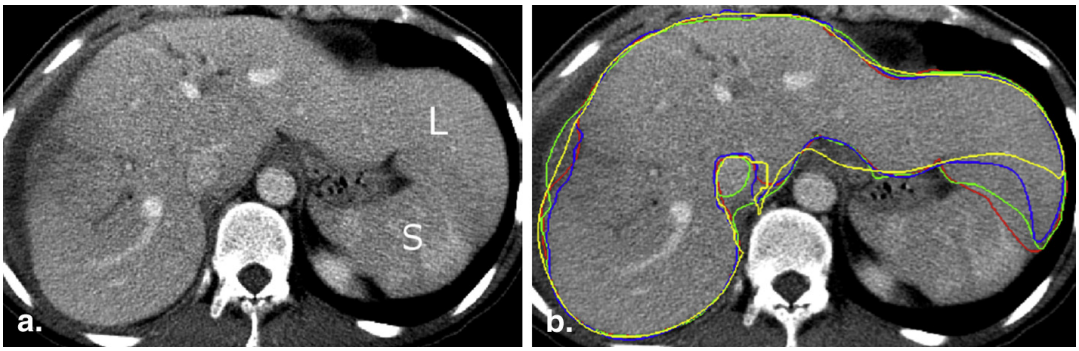

Figure 4.5: Discordant liver segmentation.

30-year-old woman with choledochal cyst. **A** and **B**, Original (**A**) and annotated (**B**) Axial CT slice demonstrating discordance between four readers using manual and semiautomated liver segmentation methods. Reader 1 manual = red tracing, Reader 2 manual = green tracing, Reader 3 semiautomated = blue tracing, Reader 4 semiautomated = yellow tracing. Discordance between readers is found at the interface between the liver (L) and the spleen (S), the liver hilum and the peripheral segment 8 liver lesion.

4.4.5 Patients with HCC

Sub-group analysis in patients with HCC (n = 4) revealed correlation between semiautomated and manual volumetry with inter-method ICC values ≥ 0.985. Repeatability studies showed intra-reader agreement for the manual method was -4 ± 17 mL with limits of agreement of -21 and 13 mL, and for the semiautomated method was 32 ± 74 mL with limits of agreement of -42 and 106 mL. The agreement between manual and semiautomated volumetry methods in this sub-group of patients was 23 ± 119 mL with limits of agreement of -96 and 142 mL.

4.4.6 Error Measures

Measures of segmentation agreement are summarized in **Table IV.IV**. All four error calculations were slightly larger for semiautomated when compared to manual methods. Volumetric overlap error was 2.9% for manual segmentation and 4.4% for semiautomated segmentation. Overall, inter-method comparisons of manual and semiautomated segmentation yielded very low error. Volumetric overlap error was 6.4 ± 1.4% (mean, standard deviation), average symmetric surface distance was 1.0 ± 0.2 mm, root mean square symmetric surface distance was 1.8 ± 0.5 mm and maximum symmetric surface distance was 17.0 ± 5.1 mm. Examples of 3D renderings with minimal and substantial surface distance error are shown in **Figure 4.6**.

Table IV.IV: Segmentation performance measures

Error Measures	Ideal Value	Intra-reader manual (R1-R1') [a]	Intra-reader semiautomated (R3-R3') [a]	Inter-method (R1-R3)
Volumetric overlap error (%)	0 %[b]	2.9 ± 0.8	4.4 ± 1.3	6.4 ± 1.4
Average symmetric surface distance (mm)	0 mm	0.4 ± 0.1	0.7 ± 0.3	1.0 ± 0.2
Root mean square symmetric surface distance (mm)	0 mm	0.9 ± 0.2	1.6 ± 0.5	1.8 ± 0.5
Maximum symmetric surface distance (mm)	0 mm	11.8 ± 4.9	17.2 ± 5.2	17.0 ± 5.1

Note: Results reported as mean ± standard deviation.

[a] R1 and R1' indicate the first and second segmentations by Reader 1 respectively. R3 and R3' indicate the first and second segmentations by Reader 3 respectively.

[b] 0% volumetric overlap error indicates perfect overlap between segmentation masks, whereas 100% volumetric overlap error indicates no overlap between segmentation masks.

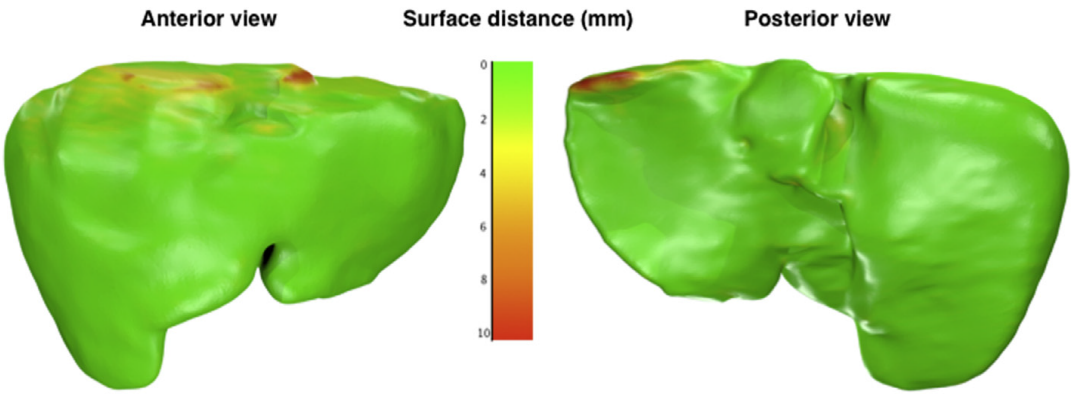

Figure 4.6: 3D surface distance error.

68-year-old woman with colorectal metastases. Anterior-posterior and posterior-anterior 3D renderings comparing surface distance error between semiautomated and manual segmentations. Areas in green represent absence of error (perfect overlap between segmentations) and areas in red represent surface distance error (in mm). Small amounts of error are observed at the liver dome and along the inferior vena cava.

4.4.7 Time

Mean interaction time was 34.3 ± 16.7 minutes per case for the manual method and 8.0 ± 1.2 minutes per case for the semiautomated method ($P < .001$)

4.5 Discussion

4.5.1 Summary of Work

This cross-sectional study evaluated the repeatability, agreement and efficiency of a semiautomated liver segmentation method by using manual segmentation as the reference standard. Overall, we found excellent correlation between semiautomated and manual segmentation volume measurements. The semiautomated method was found to have high inter-reader and intra-reader repeatability. Further, strong agreement was found between the semiautomated

and manual methods when comparing volume measurements. Finally, semiautomated liver volumetry was found to be time efficient.

Recently, an engineering competition comparing various liver segmentation algorithms on a common database of contrast-enhanced CT images was held (27). On average, interactive (i.e. semiautomated) segmentation methods that incorporated user input were found to be more accurate and reliable than fully automated methods. The larger standard deviation of automated methods was attributed to increased outlier errors (27). The study supported the use of both statistical shape information and model-based approaches to accurately represent liver structure variability (27).

In this study we evaluated a semiautomated segmentation method for CT images which did not require prior statistical information input. Our method represents a model-based approach and is a 3D extension of a technique developed for segmentation of femoral heads in biplane radiography (132). In introducing our novel method and validation framework, we address three limitations to segmentation performance described in the literature (28). First, we used a diverse surgical database to ensure that the method is reliable in pre-hepatectomy patients with a variety of hepatic pathologies. Second, we evaluated the inter-observer, intra-observer, and inter-method variability in hepatic volumes. Third, we also evaluated the quality of segmentation by using volumetric and surface error measures described in the biomedical engineering literature (27). The comparison of index and reference standard segmentation meshes permit visualization of discrepancies and provide feedback for future improvement.

4.5.2 Main Findings

In our study intra-reader, inter-reader and inter-method variability was assessed using ICC measures. Overall correlation was very good with an average ICC value of 0.995, indicating low variability in the measures. Semiautomated volumetry also achieved excellent correlation with manual volumetry (ICC ≥

0.988). A study by Suzuki et al. comparing automated and manual volumetry of living-donor livers during transplantation achieved similar results, ICC = 0.994 (17).

Bland-Altman analysis showed excellent repeatability for both manual and semiautomated methods, however the semiautomated method displayed higher bias. Unfortunately, we were not able to identify other studies which compared repeatability in this way for comparison. Semiautomated liver segmentation has inherent "problem regions" including the interface with adjacent structures, around blood vessels and in the hilum of the liver which may have lead to higher error. We attempted to limit such error by optimizing the initialization step of our segmentation method.

In our study, mean volume difference between readers for semiautomated segmentation was found to be 5 ml with limits of agreement of -93 and 103 mL in the Bland-Altman analysis. Similar inter-reader agreement was found in a study examining volumetry of resected liver specimens with achieved limits of agreement of -190 and 178 mL (135). Our narrower limits of agreement may be attributed to our readers performing segmentation on each axial slice rather than on every fourth slice as in the study by Karlo et al. (135). This potentially restricted the amount of volumetric error being interpolated to adjacent slices and led to excellent inter-observer agreement in our study.

Mean volume difference between semiautomated and manual CT-based volumetry methods was 3 mL with limits of agreement of -117 and 124 mL. These results are an improvement when compared to recently published studies which achieved limits of agreement of -230.3 and 327 mL (14), -211 and 278 mL (17) and -503 and 509 mL (13). Our narrow limits of agreement may be attributed to a variety of factors. The small degree of user feedback during manual correction of the segmentation masks likely improved the precision of semiautomated volumetry. Further, our CT-based volumetry was performed on a population of pre-hepatectomy patients rather than pre-transplant patients. The smoother liver

contours in this population, as compared to cirrhotic patients in the transplant group, may have improved agreement and precision of our method.

We anticipate that a study using the same methodology as ours on a cirrhotic population may yield less favorable results due to the more difficult segmentation inherent to nodular and dysmorphic end-stage livers. Similarly, other factors affecting hepatic parenchyma and contour, such as heterogeneous tumours, post procedural changes, or diffuse hepatic processes may also affect volume agreement. Nakayama et al. previously demonstrated that automated segmentation of damaged and deformed livers led to larger relative errors than in healthy livers (14). In our study, four patients had hepatocellular carcinoma. A review of their records revealed Child-Pugh scores between 5 and 7 (i.e. class A or B). Furthermore, imaging did not reveal dysmorphic livers except one patient who had marked segment IV atrophy. Our study was not powered to draw comparison between cirrhotic and non-cirrhotic patients by inferential statistics. Future studies on repeatability and agreement of liver segmentation may target patients with liver fibrosis or cirrhosis.

We chose to report our results according to the Bland-Altman method (134) after diligent consultation with the statistical team. Application of the Bland-Altman method for comparison between two techniques (e.g. semiautomated vs. manual segmentation) is commonly used to assess "accuracy", whereas comparison of repeated measurements (e.g. reader 1 vs. reader 1) is commonly used to assess "precision". The Bland-Altman method assumes (in order for the limits to be valid) that the error variance is constant whether expressed as a percentage or absolute value. In our article we chose to express the error as mean differences with accompanying limits of agreement for consistency with prior literature.

Overall, use of semiautomated segmentation greatly reduced the average time required for hepatic volume determination. Mean interaction time using the semiautomated method was found to be 8.0 ± 1.2 minutes per case. This is similar to recently published studies of semiautomated liver segmentation methods which

found interaction times of 20 minutes (136), 7 minutes (119) and 4.4 ± 1.9 minutes (14). Manual segmentation is often considered to be too time-consuming for clinical purposes (28). Thus, a four-fold decrease in mean interaction time is clinically relevant. Manual corrections within our interactive method remained the most-time consuming step. Improving the initialization process may reduce the need for manual correction except at liver borders, where low-contrast boundaries exist with adjacent organs (79). Wider limits of agreement were noted for semiautomated than for manual segmentation. This increased variability represents a trade-off due to faster segmentation.

To compensate for the lack of specificity of volume comparison, we incorporated four novel error metrics into our segmentation evaluation framework. These metrics apply concepts of volumetric overlap and surface distance and allow for a more robust assessment of segmentation performance. Volumetric overlap error (also known as Jaccard distance), measures the dissimilarity between two segmentation results and is defined as 1 minus the ratio of intersection and union between two segmentations; a volumetric overlap error of 0% indicates perfect overlap, which is a segmentation goal. The three remaining error metrics (average, root mean square, and maximum symmetric surface distance) are computed from the distribution of minimal distances between each surface point of the semiautomated segmentation and surface points from its corresponding manual segmentation; a value of 0 mm represents the ideal value for these 3 error metrics.

Using a variety of error metrics is preferred for broad segmentation quality evaluation (27). Our volumetric overlap error of 6.4 ± 1.4% was similar to those achieved in other studies; 5.2 ± 0.9% (119), 5.8 ± 1.4% (18)and 3.8 ± 2.2% (11). Overall, our method achieved very comparable error calculations to the best interactive segmentation methods at the MICCAI 2007 Grand Challenge (27, 119). At present time, there is no required performance specifications for error metrics, but only ideal values which are not attainable. Yet, higher values do not disqualify automated segmentation techniques as long as they are reproducible and efficient.

4.5.3 Liver Subsegmentation

Given recent surgical advances, including increases in extended hepatectomies, split-liver and living donor transplantation, establishing segmental and remnant liver volumetry is of growing importance. Subsegmentation was performed using classic vascular landmarks to divide the hepatic segments. Segmentation of the caudate lobe proved to be difficult as the boundaries were defined somewhat arbitrarily and not by vascular structures. Inter-reader correlation for segmental volumetry was found to be variable and limits of agreement were wider than those for whole liver volumetry. This can partially be explained by our choice of portal venous phase for segmentation purposes. The hepatic veins were not always clearly visible which may have increased the subjectivity in drawing the three vertical planes. In the future, alternative acquisition phases may be acquired to facilitate sub-segmentation. More reliable subsegmentation methods may also be developed based on patient-specific vascular anatomy.

4.5.4 Surgical Planning

Prior to major liver hepatectomy, the future liver remnant-to-total liver volume ratio must be calculated (77). This ratio must be > 26.5% in patients with healthy livers, > 40% in patients with high-grade steatosis, and > 50% in patients with cirrhosis (72). For this application, the level of agreement and reproducibility required is ± 5% (44). Prior to living donor liver transplantation, the liver graft-to-recipient weight ratio must be calculated. This ratio must be > 0.8% and adapted to the recipient's Child's class to avoid small-for-size syndrome (77). Although the level of agreement and reproducibility required for this application has not been specified, it is assumed to be the same as for major liver hepatectomy.

Measuring future liver remnant (FLR) volume was thought to be out of the scope of this manuscript for a variety of reasons. Our primary aim was to

accurately establish whole-liver volume as an important preliminary step before more complex segmentation procedures. Furthermore, while our dataset included a variety of liver pathologies and morphologies, not all patients specifically underwent extended hepatectomy, the usual indication for calculating FLR. Finally, determining the FLR requires a clinical judgment regarding resection margin and anticipated resection plane by a hepatobiliary surgeon. For all these reasons, calculation of this parameter was not an aim of our study because it does not lead itself to automation.

4.5.5 Segmentation Error

When visually comparing segmentation error between readers, discordance was often found at the interface between the liver and adjacent structures (stomach, diaphragm and body muscles), around blood vessels and in the hilum of the liver. Other studies have corroborated similar problem regions for liver segmentation. Heimann et al. described segmentation error at low-contrast boundaries and near tumours (27). Campadelli et al. described over-segmentation errors near the stomach and body muscles (23). Masutani et al. mentioned similar density of adjacent organs as a source of error (22). We limited such error by adjusting windowing relative to the mean liver density.

4.5.6 Limitations

Our study had some limitations. First, manual segmentation, as a reference standard, is not perfect. However, it is widely accepted in the literature and in standard clinical practice (13, 14, 17, 18, 27). Resected surgical liver volume or weight have also been described as alternative reference standards (12, 14, 15, 135). However, resected specimens can provide a false estimation of in vivo liver volume due to decreased hydrostatic pressure and blood loss from the ex vivo specimens (15, 135). Further, CT-based volumetry methods have been shown to

inaccurately estimate liver volume when compared to actual surgical resection volumes (135). These physiological variations are best avoided with the use of an in vivo reference standard such as manual segmentation.

Second, we did not perform a systematic study of segmentation robustness by varying acquisition parameters such as slice thickness and injection delays (131). Yet, the purpose of our study was to simplify workflow and shorten segmentation time while maintaining good agreement (79). Third, we did not exclude all vessels in our segmentations. Standard practice remains to exclude major vessels, but to include intrahepatic vessels in the total liver volume calculation (79).

4.5.7 Conclusion

In conclusion, our validation study suggests that a semiautomated liver segmentation method may provide high repeatability and strong agreement when compared to manual segmentation, while substantially shortening interaction time. The quality of segmentation results was confirmed by error metrics based on overlap and surface distances. Future directions include automation of segmental volumetry based on vascular anatomy (137) and adaptation of this method to MR-based liver volumetry (101).

5 Comparison of MRI and CT-based Semiautomated Liver Segmentation: a Validation Study

As first author of this manuscript, I was involved in all aspects from study design to manuscript drafting. I was an active member of the team responsible for developing the liver segmentation method by providing clinical perspective. I participated in the ethics submission, raw data collection, manual and semiautomated segmentation steps for validation and statistical analysis. I led the manuscript drafting process under the supervision of Dr. An Tang. The results of this paper were presented at the American Roentgen Ray Society annual meeting in Toronto, Canada in April 2015. This paper is currently being finalized prior to submission in a medical journal.

5.1 Abstract

Rationale and Objectives: To compare the repeatability, agreement and efficiency of MRI- and CT-based semiautomated liver segmentation for assessment of total liver volume.

Materials and Methods: This retrospective study was conducted in 31 subjects who underwent contemporaneous liver MRI and CT. Total liver volumes were segmented from contrast-enhanced 3D gradient-recalled echo MRI sequences and CT images. Semiautomated segmentation was based on variational interpolation and Laplacian mesh optimization. All segmentations were repeated after two weeks. Manual segmentation of CT images using an active contour tool was used as the reference standard. Repeatability and agreement of the methods

were evaluated with intra-class correlation coefficients (ICC) and Bland-Altman analysis. Total interaction time was recorded.

Results: Intra-reader ICC were ≥ 0.987 for MRI and ≥ 0.995 for CT. Intra-reader repeatability was 30 ± 217 ml (bias ± 1.96 SD) (95% limits of agreement: -187 to 247 ml) for MRI and -10 ± 143 ml (-153 to 133 ml) for CT. Inter-method ICC between semiautomated and manual volumetry were ≥ 0.995 for MRI and ≥ 0.986 for CT. Inter-method agreement was -14 ± 136 ml (-150 to 122 ml) for MRI and 50 ± 226 ml (-176 to 276 ml) for CT. Interaction time (mean ± SD) was significantly shorter for MRI-based semiautomated segmentation (7.2 ± 0.1 min, $p < 0.001$) and for CT-based semiautomated segmentation (6.5 ± 0.2 min, $p < 0.001$) than for CT-based manual segmentation (14.5 ± 0.4 min).

Conclusion: MRI-based semiautomated segmentation provides similar repeatability and agreement to CT-based segmentation for total liver volume.

5.2 Introduction

Accurate assessment of liver volume is fundamental in hepatic surgery prior to major hepatectomy and transplantation. Performing liver volumetry is of growing importance given recent increases in extended hepatectomies, split-liver and living-donor liver transplantations (70). Automation of liver volumetric methods has been shown to improve repeatability and accuracy while reducing processing times (18, 23, 79).

Liver segmentation has traditionally been performed on CT images due to easy accessibility, short acquisition time and high spatial resolution (14, 22, 23). However, MRI offers the advantage of simultaneous assessment of vascular and biliary anatomy and biomarkers of diffuse liver disease (fat, iron, and fibrosis) (24, 60, 65, 66). Advances in MRI techniques have prompted new indications for accurate whole liver segmentation in estimating volume-averaged biomarkers,

such as steatosis distribution maps (64, 65, 71). Furthermore, MRI minimizes the risk of radiation exposure and nephrotoxicity (24, 25).

Studies examining automated liver volumetry on MRI are limited, presumably because of increased variability and difficulty compared to CT (90). Once validated, automated liver volumetry could be integrated into a complete preoperative evaluation which includes assessment of vascular and biliary anatomy and diffuse liver disease on MRI (24, 71).

Though numerous studies have previously proposed automated segmentation methods, these have not necessarily translated to routine clinical use (27). Limitations in clinical validation, rather than lack of technical ingenuity, are thought to be the cause of this slow adaptation by the medical community (28). In order to overcome such methodological weaknesses, a validation framework for a novel automated segmentation method should include the following elements (28): use of a valid reference standard; datasets for validation which are reflective of actual clinical practice; clear metrics for measurement of segmentation precision, accuracy, efficiency and error; and comparison of metrics using effective statistical tools. We attempted to incorporate these defined elements into our validation framework.

In this article, we evaluate a novel semiautomated segmentation method which uses variational shape interpolation and a Laplacian mesh optimization framework (26). This method is compatible with both MRI and CT, which has only sparingly been previously described (89). The method does not require prior statistical input and includes mesh-based correction tools to improve precision during interactive segmentation.

5.2.1 Aims

The primary aim of our study was to compare the repeatability, agreement and efficiency of liver MRI- and CT-based semiautomated segmentation when

compared to CT-based manual segmentation. A secondary aim was to validate segmentation quality using error metrics which highlight various aspects of segmentation agreement and facilitate comparison with prior literature (27).

5.3 Materials and Methods

5.3.1 Study Design

Our institutional review board approved this retrospective, cross-sectional study. Requirements for informed consent were waived.

5.3.2 Study Subjects

Patients were included if they underwent both MRI and CT examinations within two weeks between January 2010 and March 2013 for preoperative assessment of hepatobiliary and pancreatic disease. The MRI study protocol was required to include gadolinium injection. The CT study protocol required image acquisition in portal venous phase. A total of 31 subjects (18 men, 13 women; mean age, 59 years) requiring preoperative evaluation using MRI and CT were included. These subjects had a spectrum of liver diseases. Study subjects' demographic and clinical information are summarized in **Table V.I**.

Table V.I: Subject Demographics

Characteristic	Data
Total subjects, *n* (%)	31 (100)
Sex	
Male (%)	18 (58)
Female (%)	13 (42)
Age (y), mean ± SD	59 ± 11
Weight (kg), mean ± SD	80 ± 20
Body mass index (kg/m^2), mean ± SD	28 ± 6
Pathologies, *n* (%)	
Colorectal metastases	12 (39)
Pancreatic metastases	6 (19)
Other metastasis	4 (13)
Intraductal papillary ductal neoplasm	2 (7)
Ampulloma	1 (3)
Hepatic abscess	1 (3)
Hepatocellular carcinoma	1 (3)
Klatskin tumour	1 (3)
Living donor, pre-op evaluation	1 (3)
Pancreatic pseudocyst	1 (3)
Indeterminate hepatic lesion	1 (3)

5.3.3 MRI Technique

MRI was performed with a 1.5-T unit (Discovery MR450, GE Medical Systems, Milwaukee, WI) using a 12-channel phased-array body coil. Segmentation was subsequently performed on the portal venous phase of a dynamic contrast-enhanced fat-suppressed 3-dimensional (3D) T1-weighted gradient-recalled echo (GRE) sequence (LAVA sequence). The 3D GRE sequence

parameters were: repetition time, 3.9 - 4.8 msec; echo time, 1.7 - 2.1 msec; flip angle, 12°; section thickness, 4 - 8 mm (average 5.5 mm); spacing between sections, 2.2 - 4.5 mm (average 2.7 mm); field of view, 380 mm; reconstruction matrix, 256 x 256 or 512 x 512 and parallel imaging acceleration factor, 2. A weight-adjusted dose (0.1 mmol/kg body weight) of gadobenate dimeglumine (MultiHance; Bracco Diagnostic Inc., Princeton, NJ) was administered intravenously as a bolus at a rate of 2 ml/s using a power injector (Mallinckrodt, Optistar™ Elite, St. Louis, MO), followed by saline flush of 15 ml.

5.3.4 CT Imaging Technique

CT imaging was performed with a 64-detector MDCT scanner (Brilliance 64, Philips Medical Systems, Cleveland, OH) under standard abdominal imaging protocols. The parameters were: rotation time, 0.75 seconds; detector collimation, 64 x 0.625 mm; helical pitch, 0.9; tube voltage, 120 kV; X-ray tube current: 126 - 499 mA; tube current–time product, varied based on noise index. Image reconstruction was in a 282 - 500 mm display field of view, depending on the patient's physique. Reconstruction section thickness was 2.5 mm with section gap of 2 mm. Reconstructed CT slices had a matrix size of 512 x 512 pixels with pixel spacing ranging from 0.55 - 0.78 mm. Prior to all examinations, a weight-adjusted dose of a non-ionic, low osmolar, iodinated contrast agent (375 mgI/ml Isovue; Bracco Diagnostic Inc., Princeton, NJ) was administered intravenously at a rate of 4 ml/second. All CT examinations included a portal venous phase with delay of 60 seconds.

5.3.5 Study Workflow

Liver segmentation was performed by three image analysts; two radiology residents (AG, KV; 2 and 3 years of experience respectively) and one biomedical engineering PhD candidate (GC, 3 years of experience). The image analysts were

previously trained during a CT-based liver segmentation validation study on a different data set. Two analysts independently performed semiautomated segmentation of MRI and CT images. The same analysts repeated segmentation in a random order two weeks later to prevent recall bias. A third analyst performed manual segmentation of CT images to establish the reference standard. The manual segmentation results were supervised by an abdominal radiologist (AT, 8 years of experience). Image analysts were blinded to their own segmentation results and to the results of the other readers. Interaction time was recorded for all segmentations.

5.3.6 Manual Segmentation

To establish the reference standard, axial portal venous phase CT were uploaded onto an imaging display software (SliceOmatic 4.3 Rev-11, TomoVision, Montreal, QC). Analysts manually outlined the liver using a cursor on each axial slice. This allowed for the creation of "active contours" which could be propagated to adjacent slices (91). Furthermore manual deformation of the active contours was performed for each axial image to adequately delineate the liver. Cross-sectional areas were compiled and multiplied by the slice thickness to obtain section volumes. These were added to determine the total liver volume for each patient. 3-D surface meshes created for each liver were used for visual comparison and error metric calculations. Manual segmentation of MR images was also separately performed in a similar manner. These segmentations were specifically used for error metric calculations when comparing semiautomated MRI and manual MRI surface meshes.

5.3.7 Semiautomated Segmentation

The semiautomated segmentation method was developed at the Imaging and Orthopaedics Research Laboratory (LIO, Montreal, QC) in collaboration with the

biomedical imaging team. The method is adapted from a previously validated method for CT-based liver segmentation (112), which was modified for compatibility with both MRI and CT modalities. The code was implemented in C++ using VTK (Kitware Inc., 2014, Clifton Park, NY) as a rendering external library. Contrast-enhanced MRI and CT examinations were uploaded to the segmentation program. The user (interactive) and automated (computer) tasks required for semiautomated liver segmentation are presented in **Figure 5.1**. The segmentation method consists of 3 main steps.

Initialization

In order to generate an initial shape, the user must click to position nodes around the liver contour in multi-planar views, from which a contour is interpolated. The drawn contours automatically snap onto the liver boundary using an algorithm based on image warping and minimal path segmentation (26). Generally, two contours per orthogonal plane are sufficient to generate the initial shape. An energy-minimizing implicit function (variational shape interpolation) is then applied to generate a 3D surface mesh (93, 106).

Shape Deformation

After adequate initialization of a primary liver shape, an automated optimization method is used to further refine the segmentation. Feature-matching assigns each vertex of the initial 3D surface mesh with a corresponding target point representing the most probable location of the liver boundary. This target point is determined along intensity profiles as the point of maximal intensity difference between inward (liver) and outward (non-liver) intensities. For MRI, the inward intensity is predicted for each vertex based on intensity of surrounding tissues, while for CT it is a fixed value based on estimated liver parenchymal intensity. Laplacian mesh optimization (98) is used to deform the mesh vertices

towards their matched targets on the liver boundary while ensuring a smooth local curvature.

Interactive Corrections

At times, the initial 3D surface mesh might be too distant for the intensity profile to reach the liver boundary. Additionally, adjacent structures may display similar intensity as the liver parenchyma leading to target error. For such cases, a correction tool was implemented to modify the final mesh shape. This tool allows the user to click on the surface mesh and manipulate it to the desired location. This launches a locally-constrained optimization of the mesh with relocation of adjacent vertices.

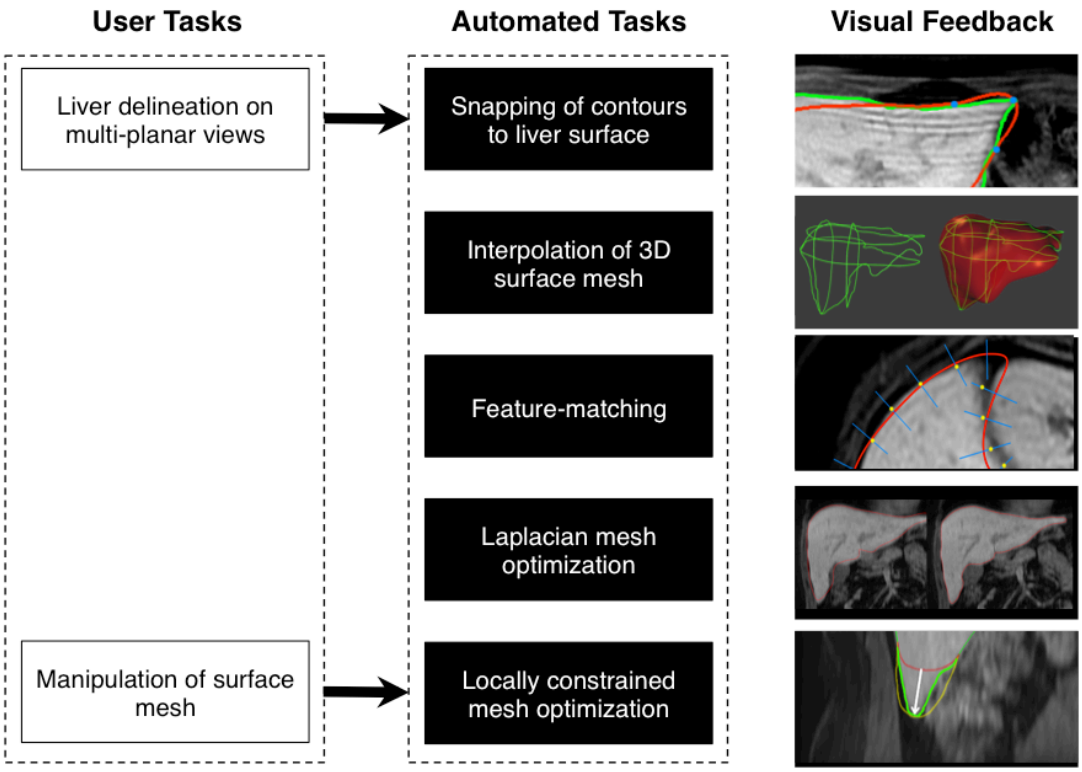

Figure 5.1: Semiautomated liver segmentation of CT and MRI images.
The user initially delineates the liver surface (2 contours per multi-planar view) from which an initial shape is defined. Variational shape interpolation is then applied to generate a 3D surface mesh. Feature-matching and Laplacian mesh optimization deform the mesh vertices towards matched targets on the actual liver boundary. The surface mesh can then be further manipulated with the aid of locally-constrained optimization.

5.3.8 Statistical Analysis

Statistical analyses was performed with SPSS software for Windows, version 21.0 (Chicago, IL). Mean whole liver volumes obtained from semiautomated segmentation of MRI and CT images were calculated by averaging the four readings for each modality. Intra-class correlation coefficients (ICC) and Bland-Altman analysis were used to determine intra-reader repeatability for semiautomated segmentation of CT and MRI images. ICC and Bland-Altman

analysis were also used to determine inter-reader and inter-method agreement, with manual CT-based segmentation being used as the reference standard for the latter. The agreement for liver volume was reported as bias ± 1.96 SD of the differences, followed by the 95% limits of agreement interval (134).

The differences between semiautomated and manually segmented surface meshes for both MRI and CT were analyzed with 4 additional error measures: volumetric overlap error, average symmetric surface distance, root mean square (RMS) symmetric surface distance and maximum symmetric surface distance. Detailed description of these error metrics can be found in a study by Heimann et al. (27). In addition, paired T-tests were used to compare total interaction time for MRI- and CT-based semiautomated segmentations with CT-based manual segmentation.

5.4 Results

5.4.1 Liver Volumes

The mean liver volume obtained from semiautomated MRI segmentations was 1831 ± 679 ml (mean ± 1.96 SD), from semiautomated CT segmentations was 1756 ± 702 ml, and from manual segmentation of CT images (reference standard) was 1817 ± 680 ml.

Detailed repeatability and agreement results are reported for both readers in **Table V.II**. To simplify the results in this section, we report the weaker (i.e. larger limits of agreement) results obtained by readers 1 or 2 below.

Table V.II: Intra-reader repeatability, inter-reader and inter-method agreement

Comparison	Readers	ICC[a]	Bland-Altman (ml)[b]
Repeatability			
Intra-reader semiautomated MRI	1 vs 1'	0.997	-19 ± 94; (-113, 75)
	2 vs 2'	0.987	30 ± 217; (-187, 247)
Intra-reader semiautomated CT	1 vs 1'	0.997	15 ± 98; (-83, 113)
	2 vs 2'	0.995	-10 ± 143; (-153, 133)
Inter-reader Agreement			
Inter-reader semiautomated MRI	1 vs 2	0.996	6 ± 123; (-117, 129)
Inter-reader semiautomated CT	1 vs 2	0.996	20 ± 125; (-105, 145)
Inter-method Agreement			
Semiautomated MRI vs. manual CT	1	0.997	-20 ± 107; (-127, 87)
	2	0.995	-14 ± 136; (-150, 122)
Semiautomated CT vs. manual CT	1	0.986	50 ± 226; (-176, 276)
	2	0.989	70 ± 202; (-132, 272)

[a] ICC = Intra-class correlation coefficient.

[b] Bland-Altman = Results reported as bias ± repeatability coefficient (1.96 SD); (95% limits of agreement interval), rounded to whole numbers.

5.4.2 Intra-reader Repeatability

The ICC were above 0.987 for MRI-based intra-reader repeatability and above 0.995 for CT-based intra-reader repeatability. Bland-Altman analysis revealed an intra-reader repeatability of 30 ± 217 ml (mean ± 1.96 SD) (95% limits of agreement: -187 to 247 ml) for MRI-based semiautomated segmentation and -10 ± 143 ml (-153 to 133 ml) for CT-based semiautomated segmentation.

5.4.3 Inter-Reader Agreement

The ICC was 0.996 for MRI- and CT-based inter-reader agreement. Bland-Altman analysis revealed an inter-reader agreement of 6 ± 123 ml (-117 to 129 ml) for MRI-based semiautomated segmentation and 20 ± 125 ml (-105 to 145 ml) for CT-based semiautomated segmentation.

5.4.4 Inter-Method Agreement

The ICC were above 0.995 for MRI-based semiautomated segmentation and above 0.986 for CT-based semiautomated segmentation when compared to manual CT. Bland-Altman analysis revealed an inter-method agreement of -14 ± 136 ml (-150 to 122 ml) for MRI-based semiautomated segmentation **(Figure 5.2)** and 50 ± 226 ml (-176 and 276 ml) for CT-based semiautomated segmentation when compared to manual CT **(Figure 5.3)**.

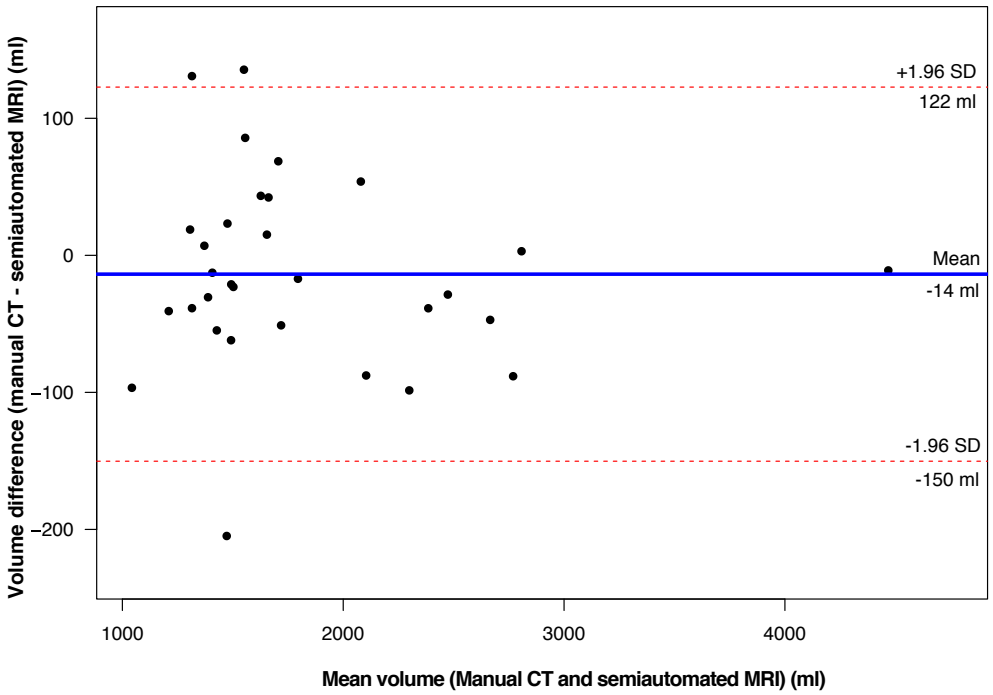

Figure 5.2: Inter-method agreement (Semiautomated MRI vs. manual CT).
Bland-Altman plot of the volume difference between MRI-based semiautomated and CT-based manual liver segmentation and their mean volume for reader 2. Mean bias demonstrated with solid line and 95% limits of agreement with dashed lines.

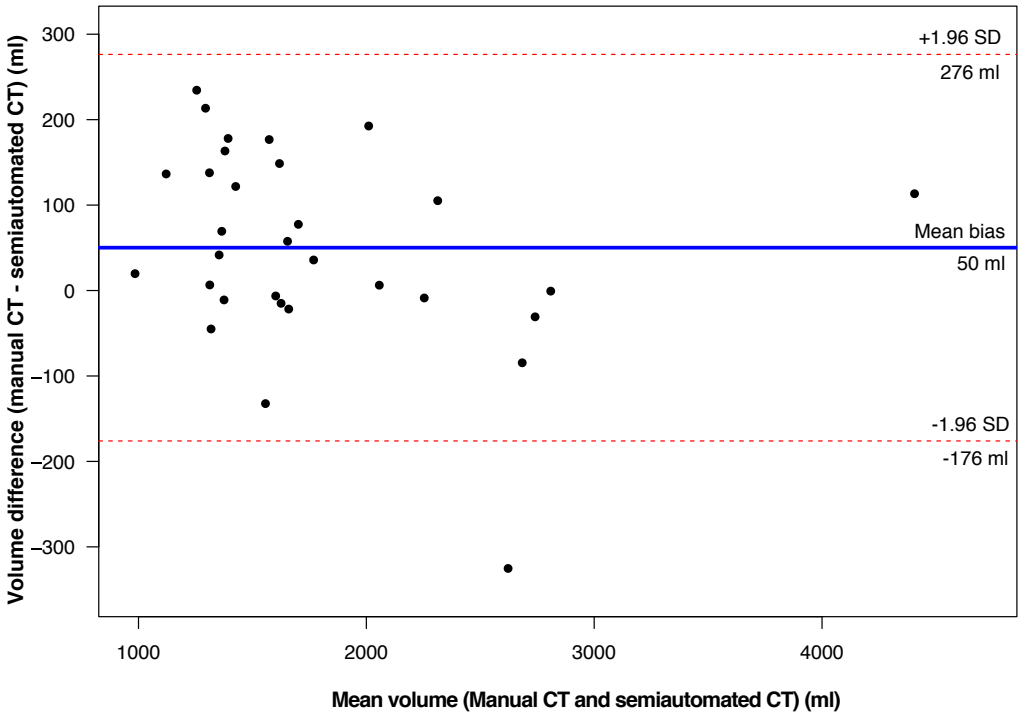

Figure 5.3: Inter-method agreement (Semiautomated CT vs. manual CT).
Bland-Altman plot of the volume difference between CT-based semiautomated and CT-based manual liver segmentation and their mean volume for reader 1. Mean bias demonstrated with solid line and 95% limits of agreement with dashed lines.

5.4.5 Clinical Examples

Examples of concordant and discordant cases displaying MRI- and CT-based semiautomated segmentations are shown in **Figures 5.4** and **5.5** respectively.

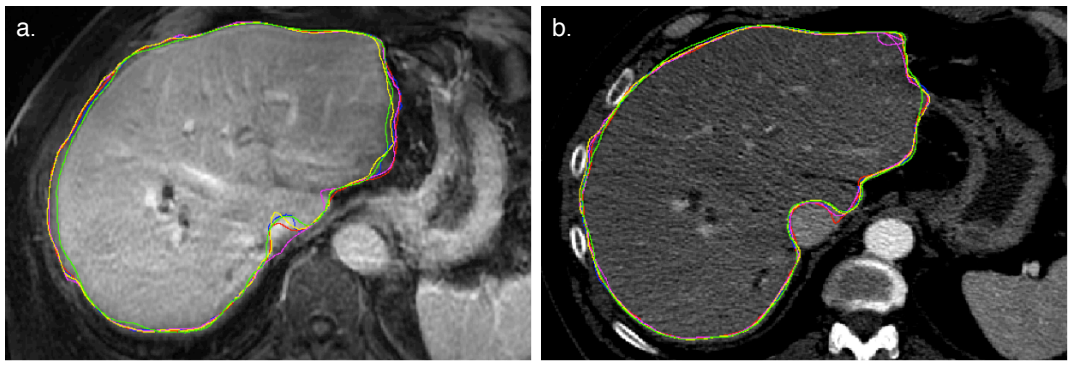

Figure 5.4: Concordant liver segmentation.

61-year-old male with a Klatskin tumour. (A) MRI and (B) CT axial images demonstrating segmentation concordance between readers using manual and semiautomated segmentation methods. Manual CT = green tracing, reader 1 semiautomated = blue tracing, reader 1' semiautomated = red tracing, reader 2 semiautomated = magenta tracing, reader 2' semiautomated = yellow tracing.

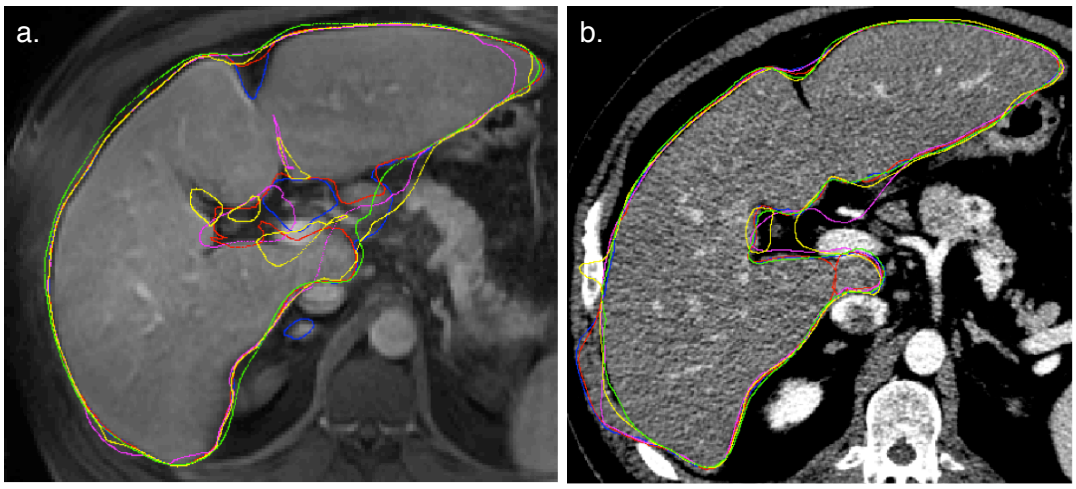

Figure 5.5: Discordant liver segmentation.

47-year-old man with pancreatic cancer metastases. (A) MRI and (B) CT axial images demonstrating segmentation discordance between readers using manual and semiautomated segmentation methods. Segmentation error on MRI and CT is noted at indistinct boundaries with adjacent organs (stomach, body muscles, vessels) and at the liver hilum. Segmentation error on MRI is also noted at convex boundaries and areas of high curvature. Manual CT = green tracing, reader 1

semiautomated = blue tracing, reader 1' semiautomated = red tracing, reader 2 semiautomated = magenta tracing, reader 2' semiautomated = yellow tracing.

5.4.6 Error Measures with MRI

Intra-reader measures of MRI-based segmentation performance were comparable between readers 1 and 2. Comparison between semiautomated MRI and manual MRI surface meshes for reader 2 revealed volumetric overlap error of 11.6 ± 3.4% (mean ± standard deviation), average symmetric surface distance was 2.3 ± 0.6 mm, root mean square symmetric surface distance was 28.0 ± 10.4 mm and maximum symmetric surface distance was 3.8 ± 1.2 mm **(Table V.III)**.

Table V.III: Segmentation performance measures for MRI

Error Measures	Ideal Value	Intra-reader: semiautomated MRI		Inter-method: semiautomated MRI vs manual MRI	
		R1 vs R1'[a]	R2 vs. R2'	R1	R2
Volumetric overlap error (%)	0 %[b]	5.6 ± 1.9	8.9 ± 5.2	9.2 ± 1.9	11.6 ± 3.4
Average symmetric surface distance (mm)	0 mm	1.2 ± 0.4	1.7 ± 0.9	1.9 ± 0.4	2.3 ± 0.6
Root mean square symmetric surface distance (mm)	0 mm	21.3 ± 6.4	25.0 ± 10.6	24.9 ± 9.9	28.0 ± 10.4
Maximum symmetric surface distance (mm)	0 mm	2.5 ± 0.8	3.1 ± 1.4	3.2 ± 0.9	3.8 ± 1.2

Note: Results reported as mean ± standard deviation.

[a] R1 and R1' indicate the first and second segmentations by reader 1 respectively.

[b] 0% volumetric overlap error indicates perfect overlap between segmentation masks, whereas 100% volumetric overlap error indicates no overlap between segmentation masks.

5.4.7 Error Measures with CT

Intra-reader measures of CT-based segmentation performance were comparable between readers 1 and 2. Comparison between semiautomated CT and manual CT surface meshes for reader 1 revealed volumetric overlap error of 9.2 ± 2.5% (mean ± standard deviation), average symmetric surface distance of 1.7 ± 0.4 mm, root mean square symmetric surface distance of 24.9 ± 6.9 mm and maximum symmetric surface distance of 3.0 ± 0.9 mm **(Table V.IV)**.

Table V.IV: Segmentation performance measures for CT

Error Measures	**Ideal Value**	**Intra-reader: semiautomated CT**		**Inter-method: semiautomated CT vs manual CT**	
		R1 vs R1'[a]	**R2 vs. R2'**	**R1**	**R2**
Volumetric overlap error (%)	0 %[b]	5.8 ± 2.7	5.8 ± 1.7	9.2 ± 2.5	9.1 ± 2.1
Average symmetric surface distance (mm)	0 mm	1.2 ± 0.5	1.0 ± 0.3	1.7 ± 0.4	1.8 ± 0.4
Root mean square symmetric surface distance (mm)	0 mm	27.5 ± 15.7	23.4 ± 7.4	24.9 ± 6.9	27.7 ± 14.1
Maximum symmetric surface distance (mm)	0 mm	2.8 ± 1.6	2.2 ± 0.7	3.0 ± 0.9	3.2 ± 1.4

Note: Results reported as mean ± standard deviation.

[a] R1 and R1' indicate the first and second segmentations by reader 1 respectively.

[b] 0% volumetric overlap error indicates perfect overlap between segmentation masks, whereas 100% volumetric overlap error indicates no overlap between segmentation masks.

5.4.8 Time

Interaction time (mean ± SD) per case was significantly shorter for MRI-based semiautomated segmentation (7.2 ± 0.1 min, $p < 0.001$) and for CT-based semiautomated segmentation (6.5 ± 0.2 min, $p < 0.001$) than for CT-based manual segmentation (14.5 ± 0.4 min).

5.5 Discussion

5.5.1 Summary of Work

This retrospective study evaluated the repeatability, agreement and efficiency of MRI- and CT-based semiautomated segmentation, using CT-based manual segmentation as the reference standard method. A strength of our study lies in the paired comparison of two imaging modalities, while using the same independent reference standard. Our choice of a semiautomated liver segmentation method was supported by recent studies which found interactive methods to be generally more accurate and reliable than fully automated methods (27). Segmentation was easily customized for MRI and CT using a varying feature-matching strategy, demonstrating the multi-modality versatility of our method.

5.5.2 Main Findings

Overall, semiautomated volume measurements for both MRI and CT strongly correlated with volumes obtained by manual segmentation. MRI-based and CT-based semiautomated volumetry were highly repeatable and showed strong

agreement with the manual method. Intra-reader repeatability for MRI-based semiautomated segmentation was comparable to the results for CT. However, Bland-Altman analysis showed slightly higher repeatability coefficients compared to previous studies evaluating automated segmentation methods with two readers. Mazonakis et al. (83) examined 38 consecutive patients referred for MRI examination and found repeatability coefficients of 51.6 and 68.2 ml, while Hermoye et al. (79) studied 18 liver donors and found repeatability coefficients of 52 and 64 ml. Our study examined only pathological livers which may explain increased variability in the repeatability calculation.

Our study showed superior ICC values for MRI-based inter-method agreement compared to prior studies: 0.98 (85, 89, 90) and 0.76 - 0.93 (88). Further, our limits of agreement were similar to those obtained in recent studies: -108 to 91 ml (83), -163 to 134 ml (90), and -278 to 204 ml (88).

Inter-method agreement between CT-based semiautomated segmentation and manual segmentation compared favorably to recently published studies. Previous studies have shown ICC values for CT-based semiautomated segmentation from 0.94 to 0.994 (17, 89) and limits of agreement of -117 to 124 ml (112), -230.3 to 327 ml (14), -211 to 278 ml (17), and -503 to 509 ml (13).

Segmentation quality was further evaluated with volumetric and surface error measures which have previously been used in the setting of segmentation evaluation frameworks (112, 138). The comparison of meshes obtained from semiautomated and manual methods also aided in the direct visualization of segmentation discrepancies. In order to adequately evaluate MR segmentation meshes, we also performed manual MR segmentation. This was required as inter-modality comparison of meshes (i.e. semiautomated MR and manual CT) are not possible due to inherent differences in image acquisition such as variable breath-holds and elastic liver deformation. Such a comparison would result in misregistration of meshes and artificial elevation of surface error measures.

For MRI, the comparison between semiautomated and manual surface meshes revealed a volumetric overlap error similar to CT: 11.6 ± 3.4%. Our results were comparable to another MRI-based automated segmentation study which achieved volumetric overlap error of 11.2%, average symmetric surface distance of 2.2 mm, and maximum symmetric surface distance of 34 mm (103).

Semiautomated segmentation significantly reduced the interaction time required for determination of liver volume. Recently published studies have described semiautomated segmentation times ranging from 8 ± 2 min to 13.3 ± 4.5 min for MRI-based methods (71, 85) and from 4.4 ± 1.9 min to 8.0 ± 1.2 min for CT-based methods (14, 112, 139).

5.5.3 Segmentation Error

Segmentation errors on MRI were noted at similar locations to CT: primarily at the liver interface with adjacent structures (muscles, diaphragm, spleen, stomach), at the liver hilum, adjacent to tumours and near blood vessels. In addition, areas of convex and concave boundaries and high curvature (such as liver dome) contributed significantly to segmentation error. Under-segmentation on MRI occurred at low-contrast liver boundaries and areas of inhomogeneous density whereas over-segmentation usually occurred at organs abutting the liver, as noted by Huynh et al. (90). Motion, pulsation and partial volume artifacts have also been shown to impede segmentation accuracy.

5.5.4 Limitations

Our study had certain limitations. First, our choice of manual CT segmentation as the reference standard for validating a MRI-based semiautomated method had not previously been described. Other studies validating automated MRI methods have traditionally relied on manual MRI segmentation as the reference standard. We opted for an independent reference standard in order to

validate both MRI- and CT-based semiautomated segmentation as our method is compatible with both. This common standard promotes head-to-head comparison of automated segmentation accuracy between MRI and CT, which previously was not addressed. Manual CT segmentation has been used as the reference standard in numerous other similar studies. Resected surgical volume or weight may also have been alternate reference standards. However, in vivo liver volume may be falsely estimated due to blood loss and changes in hydrostatic pressure following surgical resection (15, 107, 135), thus making manual segmentation a more reliable choice.

Second, our validation scheme utilized similar MRI acquisition parameters as we did not perform a systematic study of segmentation robustness. As in previous studies, a 3D T1-weighted GRE sequence was used for MRI-based segmentation (107, 135). The portal venous phase was chosen as it maximizes contrast between the liver and adjacent structures (17).

Third, subsegmentation based on patient-specific vascular anatomy was not performed. Further, as in previous studies, intrahepatic vessels were included in the liver volume assessment though major vessels were excluded (79). Future research directions may include automated vascular sub-segmentation to accurately determine segmental liver volumetry.

5.5.5 Conclusion

In conclusion, our validation study suggests that a semiautomated liver segmentation method compatible with both MRI and CT can provide strong agreement and repeatability when compared to manual segmentation, while shortening interaction time. Given recent advances in MRI-based biomarkers of chronic liver disease, accurate estimation of liver volume using MRI is of significance. Automated volumetry could also be integrated into a complete MRI-based preoperative evaluation to assess vascular and biliary anatomy and liver quality. Future studies may validate alternative MRI sequences for liver

volumetry, particularly fat quantification sequences and study vascular sub-segmentation.

6 Conclusion

6.1 Future Work

Vascular subsegmentation

The Couinaud classification scheme is a simple system to divide the liver into subsegments based on expected vascular anatomy. This scheme has been criticized as it provides an anatomic estimation and is not tailored to patient-specific anatomy. Future trends include liver subsegmentation based on supplying (i.e. hepatic artery, portal vein) or draining (i.e. hepatic vein) vasculature. This will require segmentation of all vascular structures as well as the hepatic parenchyma itself **(Figure 6.1)**. Patient-specific subsegmentation, which takes into account various anatomic variants, will provide surgeons with vital anatomic information prior to complex hepatic resections and living-donor transplantations.

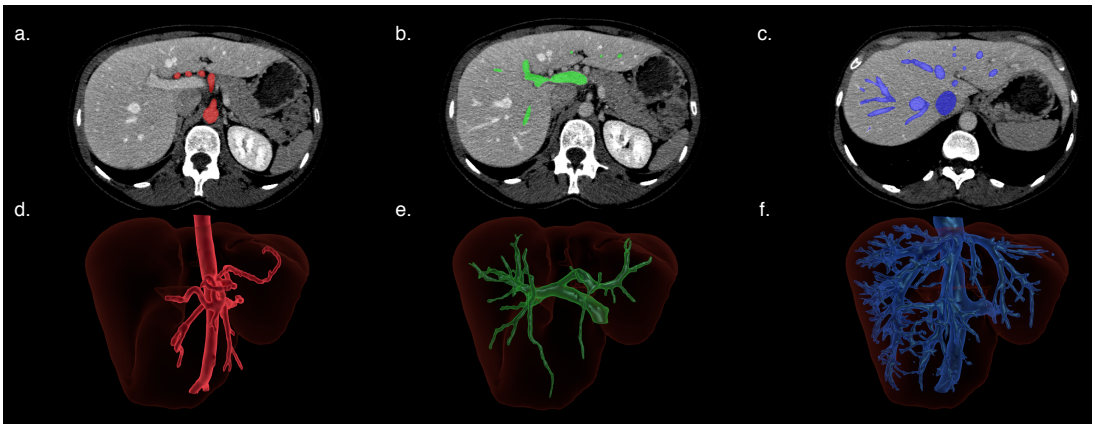

Figure 6.1: Liver subsegmentation according to vascular anatomy.
Axial contrast-enhanced CT scan shows the segmented (a) arterial, (b) portal venous, and (c) hepatic venous structures. 3D rendering in the same patient shows the corresponding segmented (d) arterial, (e) portal venous vascular structures, and (f) hepatic venous structures (8).

Virtual Surgical planning

Thought private companies offer a variety of liver segmentation and surgical planning services, they are often expensive and time-consuming to obtain. The related cost makes it difficult for each patient to undergo a virtual surgical planning procedure prior to hepatecomy or transplantation. Future trends in liver segmentation include in-house solutions for surgical planning. Accurate parenchymal and vascular segmentation must be combined with 3D modeling to provide such solutions **(Figure 6.2)**.

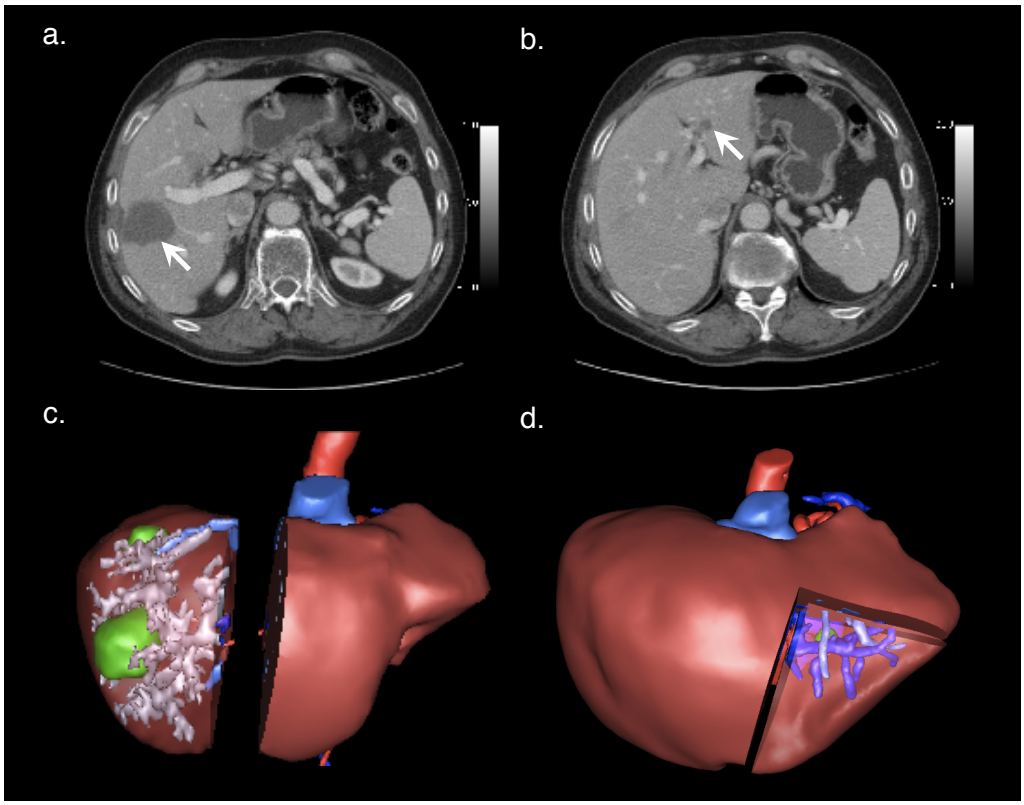

Figure 6.2: Virtual Surgical Planning.

Tumour visualization: (a) Axial enhanced CT image shows a right liver metastasis centered in segment V (arrow). The patient also had a metastasis involving segment VII (not shown). (b) Axial enhanced CT image shows a left liver metastasis in segment III (arrow). **Surgical simulation and volume planning:** (c) 3D-rendering image shows surgical planning for complete right hepatectomy. (d) 3D-rendering image shows surgical planning for segmentectomy of segment III. Residual hepatic liver volume after both procedures was estimated to be 27%. Right portal embolization was thus performed before right hepatectomy. Images 6.2c and d courtesy of Dr. Franck Vandenbroucke-Menu (8).

6.2 Closing Words

Liver disease is an important public health concern with substantial morbidity and mortality associated with disease progression. There is a significant need for better tools to diagnose and monitor biomarkers of liver pathology. Our

segmentation software represents a non-invasive, imaging-based method to procure vital information regarding 3D liver shape and volume.

The research presented in this dissertation is part of a larger research program involving MR-based quantification of liver biomarkers coordinated by my research director, Dr. An Tang. The development and validation of our segmentation method contributes to this program by providing a tool for accurate and reliable liver volume assessment.

If validated for clinical usage, this method could be integrated with other MRI-based biomarkers of disease (i.e., fat, iron, fibrosis, inflammation) to determine pathology density throughout the entire liver volume. Conversely liver biopsy, the current reference standard, provides sampling of an extremely small portion of the liver and is an invasive procedure.

In the development of our method, we opted for a model-based approach which used implicit modeling for shape initialization combined with a shape deformation scheme based on Laplacian mesh optimization. A primary goal was to develop a method compatible with both CT and MRI images as this has rarely been described in the medical literature. Further, we aspired to overcome the need for training data while maintaining accuracy and repeatability on a wide range of pathological livers.

Results from our cross-sectional studies suggest that we were able to achieve these stated goals. In doing so, we demonstrated that semiautomated liver segmentation represents a feasible alternative to volumetric methods currently employed in the clinical setting.

My hope is that this dissertation provides the reader with a more sophisticated understanding of automated liver volumetry from the clinical perspective. I look forward to the day when I am able to quickly evaluate liver volume and other biomarkers of liver disease for each patient while reviewing their CT or MRI examinations.

Bibliography

Publications by our research team are in bold characters

1. Caldwell RS. Hesiod's Theogony. Cambridge, MA: Focus Information Group 1987;1987:59.
2. Chen TS, Chen PS. The myth of Prometheus and the liver. Journal of the Royal Society of Medicine. 1994;87(12):754-5.
3. West ML. Hesiod: Theogony. Oxford: Clarendon Press. 1966.
4. Michalopoulos GK. Liver regeneration. Journal of cellular physiology. 2007;213(2):286-300. Epub 2007/06/15.
5. Boyer TD, Manns MP, Sanyal AJ, Zakim D. Zakim and Boyer's Hepatology: A Textbook of Liver Disease: Saunders/Elsevier; 2012.
6. Couinaud C. Le Foie: études anatomiques et chirurgicales. Paris: Masson. 1957.
7. Fasel JH, Schenk A. Concepts for Liver Segment Classification: Neither Old Ones nor New Ones, but a Comprehensive One. Journal of clinical imaging science. 2013;3:48. Epub 2013/11/15.
8. **Gotra A, Chartrand G, Vu K, Vandenbroucke-Menu F, Kauffmann C, Gallix B, et al., editors. Liver Segmentation: A Primer for Radiologists. Radiological Society of North America 2014 Scientific Assembly and Annual Meeting; 2014; Chicago, IL.**
9. Zoli M, Cordiani MR, Marchesini G, Iervese T, Labate AM, Bonazzi C, et al. Prognostic indicators in compensated cirrhosis. The American journal of gastroenterology. 1991;86(10):1508-13. Epub 1991/10/01.
10. Sekiyama K, Yoshiba M, Inoue K, Sugata F. Prognostic value of hepatic volumetry in fulminant hepatic failure. Digestive diseases and sciences. 1994;39(2):240-4. Epub 1994/02/01.

11. Linguraru MG, Sandberg JK, Jones EC, Petrick N, Summers RM. Assessing hepatomegaly: automated volumetric analysis of the liver. Academic radiology. 2012;19(5):588-98. Epub 2012/03/01.
12. Yamanaka J, Saito S, Fujimoto J. Impact of preoperative planning using virtual segmental volumetry on liver resection for hepatocellular carcinoma. World journal of surgery. 2007;31(6):1249-55.
13. Luciani A, Rusko L, Baranes L, Pichon E, Loze B, Deux JF, et al. Automated liver volumetry in orthotopic liver transplantation using multiphase acquisitions on MDCT. Ajr. 2012;198(6):W568-74. Epub 2012/05/25.
14. Nakayama Y, Li Q, Katsuragawa S, Ikeda R, Hiai Y, Awai K, et al. Automated hepatic volumetry for living related liver transplantation at multisection CT. Radiology. 2006;240(3):743-8. Epub 2006/07/22.
15. Lemke AJ, Brinkmann MJ, Schott T, Niehues SM, Settmacher U, Neuhaus P, et al. Living donor right liver lobes: preoperative CT volumetric measurement for calculation of intraoperative weight and volume. Radiology. 2006;240(3):736-42.
16. Kamel IR, Kruskal JB, Pomfret EA, Keogan MT, Warmbrand G, Raptopoulos V. Impact of multidetector CT on donor selection and surgical planning before living adult right lobe liver transplantation. Ajr. 2001;176(1):193-200.
17. Suzuki K, Epstein ML, Kohlbrenner R, Garg S, Hori M, Oto A, et al. Quantitative radiology: automated CT liver volumetry compared with interactive volumetry and manual volumetry. Ajr. 2011;197(4):W706-12. Epub 2011/09/24.
18. Massoptier L, Casciaro S. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. European radiology. 2008;18(8):1658-65. Epub 2008/03/29.
19. Castell DO, O'Brien KD, Muench H, Chalmers TC. Eastimation of liver size by percussion in normal individuals. Annals of internal medicine. 1969;70(6):1183-9. Epub 1969/06/01.

20. Castell DO, Frank BB. Abdominal examination: role of percussion and auscultation. Postgraduate medicine. 1977;62(6):131-4. Epub 1977/12/01.
21. Heymsfield SB, Fulenwider T, Nordlinger B, Barlow R, Sones P, Kutner M. Accurate measurement of liver, kidney, and spleen volume and mass by computerized axial tomography. Annals of internal medicine. 1979;90(2):185-7. Epub 1979/02/01.
22. Masutani Y, Uozumi K, Akahane M, Ohtomo K. Liver CT image processing: a short introduction of the technical elements. European journal of radiology. 2006;58(2):246-51.
23. Campadelli P, Casiraghi E, Esposito A. Liver segmentation from computed tomography scans: a survey and a new algorithm. Artif Intell Med. 2009;45(2-3):185-96. Epub 2008/12/09.
24. Fulcher AS, Szucs RA, Bassignani MJ, Marcos A. Right lobe living donor liver transplantation: preoperative evaluation of the donor with MR imaging. Ajr. 2001;176(6):1483-91.
25. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. The New England journal of medicine. 2007;357(22):2277-84. Epub 2007/11/30.
26. Chartrand G, Cresson T, Chav R, Gotra A, Tang A, De Guise J, editors. Semi-automated liver CT segmentation using Laplacian meshes. 2014 IEEE International Symposium on Biomedical Imaging; 2013; Beiing, China.
27. Heimann T, van Ginneken B, Styner MA, Arzhaeva Y, Aurich V, Bauer C, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans Med Imaging. 2009;28(8):1251-65. Epub 2009/02/13.
28. Udupa JK, Leblanc VR, Zhuge Y, Imielinska C, Schmidt H, Currie LM, et al. A framework for evaluating image segmentation algorithms. Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society. 2006;30(2):75-87. Epub 2006/04/06.

29. Sherman M, Bilodeau M, Cooper C, Mackie D, Depew W, Villeneuve J-P, et al. Liver Disease in Canada: A Crisis in the Making. An Assessment of Liver Disease in Canada. 2013 March 2013. Report No.
30. Chaudhary RK, Nicholls ES, Kennedy DA. Prevalence of hepatitis B markers in Indochinese refugees. Canadian Medical Association journal. 1981;125(11):1243-6. Epub 1981/12/01.
31. Wong WW, Minuk GY. A cross-sectional seroepidemiologic survey of chronic hepatitis B virus infections in Southeast Asian immigrants residing in a Canadian urban centre. Clinical and investigative medicine Medecine clinique et experimentale. 1994;17(5):443-7. Epub 1994/10/01.
32. Strickland GT, El-Kamary SS, Klenerman P, Nicosia A. Hepatitis C vaccine: supply and demand. The Lancet Infectious diseases. 2008;8(6):379-86. Epub 2008/05/27.
33. Brunt EM. Pathology of nonalcoholic steatohepatitis. Hepatology research : the official journal of the Japan Society of Hepatology. 2005;33(2):68-71. Epub 2005/10/11.
34. Brunt EM. Nonalcoholic steatohepatitis. Seminars in liver disease. 2004;24(1):3-20. Epub 2004/04/16.
35. Sheth SG, Gordon FD, Chopra S. Nonalcoholic steatohepatitis. Annals of internal medicine. 1997;126(2):137-45. Epub 1997/01/15.
36. Preiss D, Sattar N. Non-alcoholic fatty liver disease: an overview of prevalence, diagnosis, pathogenesis and treatment considerations. Clin Sci (Lond). 2008;115(5):141-50. Epub 2008/07/30.
37. Gupta AA, Kim DC, Krinsky GA, Lee VS. CT and MRI of cirrhosis and its mimics. Ajr. 2004;183(6):1595-601. Epub 2004/11/18.
38. Abbas AK, Cotran SC, Kumar V, Robbins SL, Fausto N. Robbins & Cotran Pathologic Basis of Disease: Saunders; 7 edition; 2004.
39. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. International journal of cancer Journal international du cancer. 2001;94(2):153-6. Epub 2001/10/23.

40. Mitchell DG, Bruix J, Sherman M, Sirlin CB. LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. Hepatology. 2015;61(3):1056-65. Epub 2014/07/22.
41. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology. 2011;53(3):1020-2. Epub 2011/03/05.
42. Namasivayam S, Martin DR, Saini S. Imaging of liver metastases: MRI. Cancer imaging : the official publication of the International Cancer Imaging Society. 2007;7:2-9. Epub 2007/02/13.
43. Doherty GM, Way LW. Current Surgical Diagnosis & Treatment: Lange Medical Books/McGraw-Hill; 2006.
44. Abdalla EK, Adam R, Bilchik AJ, Jaeck D, Vauthey JN, Mahvi D. Improving resectability of hepatic colorectal metastases: expert consensus statement. Annals of surgical oncology. 2006;13(10):1271-80. Epub 2006/09/07.
45. Geoghegan JG, Scheele J. Treatment of colorectal liver metastases. Br J Surg. 1999;86(2):158-69. Epub 1999/04/01.
46. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clinical pharmacology and therapeutics. 2001;69(3):89-95. Epub 2001/03/10.
47. Ramachandran U, Alurkar V, Thaplia A. Pattern of cardiac diseases in children in Pokhara, Nepal. Kathmandu Univ Med J (KUMJ). 2006;4(2):222-7. Epub 2008/07/08.
48. Manolio T. Novel risk markers and clinical practice. The New England journal of medicine. 2003;349(17):1587-9. Epub 2003/10/24.
49. Aithal GP, Guha N, Fallowfield J, Castera L, Jackson AP. Biomarkers in liver disease: emerging methods and potential applications. International journal of hepatology. 2012;2012:437508. Epub 2012/12/05.
50. Okazaki H, Ito K, Fujita T, Koike S, Takano K, Matsunaga N. Discrimination of alcoholic from virus-induced cirrhosis on MR imaging. Ajr. 2000;175(6):1677-81. Epub 2000/11/25.

51. Zhou X, Kitagawa T, Hara T, Fujita H, Zhang X, Yokoyama R, et al. Constructing a probabilistic model for automated liver region segmentation using non-contrast X-ray torso CT images. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv. 2006;9(Pt 2):856-63.

52. Bora A, Alptekin C, Yavuz A, Batur A, Akdemir Z, Berkoz M. Assessment of liver volume with computed tomography and comparison of findings with ultrasonography. Abdom Imaging. 2014;39(6):1153-61. Epub 2014/04/30.

53. Hagan MT, Sayuk GS, Lisker-Melman M, Korenblat KM, Kerr TA, Chapman WC, et al. Liver volume in the cirrhotic patient: does size matter? Digestive diseases and sciences. 2014;59(4):886-91. Epub 2014/02/08.

54. Fernandez-Salazar L, Velayos B, Aller R, Lozano F, Garrote JA, Gonzalez JM. Percutaneous liver biopsy: patients' point of view. Scand J Gastroenterol. 2011;46(6):727-31. Epub 2011/03/04.

55. Bravo AA, Sheth SG, Chopra S. Liver biopsy. The New England journal of medicine. 2001;344(7):495-500. Epub 2001/02/15.

56. Joy D, Thava VR, Scott BB. Diagnosis of fatty liver disease: is biopsy necessary? European journal of gastroenterology & hepatology. 2003;15(5):539-43. Epub 2003/04/19.

57. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. The American journal of gastroenterology. 2002;97(10):2614-8. Epub 2002/10/19.

58. Longo R, Pollesello P, Ricci C, Masutti F, Kvam BJ, Bercich L, et al. Proton MR spectroscopy in quantitative in vivo determination of fat content in human liver steatosis. J Magn Reson Imaging. 1995;5(3):281-5. Epub 1995/05/01.

59. Szczepaniak LS, Babcock EE, Schick F, Dobbins RL, Garg A, Burns DK, et al. Measurement of intracellular triglyceride stores by H spectroscopy:

validation in vivo. The American journal of physiology. 1999;276(5 Pt 1):E977-89. Epub 1999/05/18.

60. Reeder SB, Cruite I, Hamilton G, Sirlin CB. Quantitative assessment of liver fat with magnetic resonance imaging and spectroscopy. J Magn Reson Imaging. 2011;34(4):729-49. Epub 2011/09/20.

61. Yokoo T, Bydder M, Hamilton G, Middleton MS, Gamst AC, Wolfson T, et al. Nonalcoholic fatty liver disease: diagnostic and fat-grading accuracy of low-flip-angle multiecho gradient-recalled-echo MR imaging at 1.5 T. Radiology. 2009;251(1):67-76. Epub 2009/02/18.

62. Yokoo T, Shiehmorteza M, Hamilton G, Wolfson T, Schroeder ME, Middleton MS, et al. Estimation of hepatic proton-density fat fraction by using MR imaging at 3.0 T. Radiology. 2011;258(3):749-59. Epub 2011/01/08.

63. Dixon WT. Simple proton spectroscopic imaging. Radiology. 1984;153(1):189-94. Epub 1984/10/01.

64. Le TA, Chen J, Changchien C, Peterson MR, Kono Y, Patton H, et al. Effect of colesevelam on liver fat quantified by magnetic resonance in nonalcoholic steatohepatitis: a randomized controlled trial. Hepatology. 2012;56(3):922-32. Epub 2012/03/21.

65. Tang A, Chen J, Le TA, Changchien C, Hamilton G, Middleton MS, et al. Cross-sectional and longitudinal evaluation of liver volume and total liver fat burden in adults with nonalcoholic steatohepatitis. Abdom Imaging. 2015;40(1):26-37. Epub 2014/07/13.

66. Hernando D, Levin YS, Sirlin CB, Reeder SB. Quantification of liver iron with MRI: state of the art and remaining challenges. J Magn Reson Imaging. 2014;40(5):1003-21. Epub 2014/03/04.

67. Imamura H, Seyama Y, Kokudo N, Aoki T, Sano K, Minagawa M, et al. Single and multiple resections of multiple hepatic metastases of colorectal origin. Surgery. 2004;135(5):508-17. Epub 2004/05/01.

68. Poon RT, Fan ST, Lo CM, Ng IO, Liu CL, Lam CM, et al. Improving survival results after resection of hepatocellular carcinoma: a prospective

study of 377 patients over 10 years. Annals of surgery. 2001;234(1):63-70. Epub 2001/06/23.

69. Pang YY. The Brisbane 2000 terminology of liver anatomy and resections. HPB 2000; 2:333-39. HPB : the official journal of the International Hepato Pancreato Biliary Association. 2002;4(2):99; author reply -100. Epub 2008/03/12.

70. Ferrero A, Vigano L, Polastri R, Muratore A, Eminefendic H, Regge D, et al. Postoperative liver dysfunction and future remnant liver: where is the limit? Results of a prospective study. World J Surg. 2007;31(8):1643-51. Epub 2007/06/07.

71. d'Assignies G, Kauffmann C, Boulanger Y, Bilodeau M, Vilgrain V, Soulez G, et al. Simultaneous assessment of liver volume and whole liver fat content: a step towards one-stop shop preoperative MRI protocol. European radiology. 2011;21(2):301-9. Epub 2010/09/04.

72. Lim MC, Tan CH, Cai J, Zheng J, Kow AW. CT volumetry of the liver: where does it stand in clinical practice? Clin Radiol. 2014;69(9):887-95. Epub 2014/05/16.

73. Vauthey JN, Chaoui A, Do KA, Bilimoria MM, Fenstermacher MJ, Charnsangavej C, et al. Standardized measurement of the future liver remnant prior to extended liver resection: methodology and clinical associations. Surgery. 2000;127(5):512-9. Epub 2000/05/20.

74. Low HC, Da Costa M, Prabhakaran K, Kaur M, Wee A, Lim SG, et al. Impact of new legislation on presumed consent on organ donation on liver transplant in Singapore: a preliminary analysis. Transplantation. 2006;82(9):1234-7. Epub 2006/11/15.

75. Busuttil RW, Goss JA. Split liver transplantation. Annals of surgery. 1999;229(3):313-21. Epub 1999/03/17.

76. Lo CM, Fan ST, Liu CL, Wei WI, Lo RJ, Lai CL, et al. Adult-to-adult living donor liver transplantation using extended right lobe grafts. Annals of surgery. 1997;226(3):261-9; discussion 9-70. Epub 1997/10/27.

77. Ben-Haim M, Emre S, Fishbein TM, Sheiner PA, Bodian CA, Kim-Schluger L, et al. Critical graft size in adult-to-adult living donor liver transplantation: impact of the recipient's disease. Liver Transpl. 2001;7(11):948-53. Epub 2001/11/08.

78. Kiuchi T, Kasahara M, Uryuhara K, Inomata Y, Uemoto S, Asonuma K, et al. Impact of graft size mismatching on graft prognosis in liver transplantation from living donors. Transplantation. 1999;67(2):321-7. Epub 1999/02/12.

79. Hermoye L, Laamari-Azjal I, Cao Z, Annet L, Lerut J, Dawant BM, et al. Liver Segmentation in Living Liver Transplant Donors: Comparison of Semiautomatic and Manual Methods. Radiology. 2005;234(1):171-8.

80. Kiuchi T, Tanaka K, Ito T, Oike F, Ogura Y, Fujimoto Y, et al. Small-for-size graft in living donor liver transplantation: how far should we go? Liver Transpl. 2003;9(9):S29-35. Epub 2003/08/28.

81. Lemke AJ, Hosten N, Neumann K, Muller B, Neuhaus P, Felix R, et al. [CT volumetry of the liver before transplantation]. RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin. 1997;166(1):18-23. Epub 1997/01/01. CT-Volumetrie der Leber vor Transplantation.

82. Frericks BB, Kiene T, Stamm G, Shin H, Galanski M. CT-based liver volumetry in a porcine model: impact on clinical volumetry prior to living donated liver transplantation. RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin. 2004;176(2):252-7. Epub 2004/02/12. CT-basierte Lebervolumetrie im Tiermodell: Bedeutung fur die klinische Volumetrie im Rahmen der Leberlebendspende.

83. Mazonakis M, Damilakis J, Maris T, Prassopoulos P, Gourtsoyiannis N. Comparison of two volumetric techniques for estimating liver volume using magnetic resonance imaging. J Magn Reson Imaging. 2002;15(5):557-63. Epub 2002/05/09.

84. Hermoye L, Laamari-Azjal I, Cao Z, Annet L, Lerut J, Dawant BM, et al. Liver segmentation in living liver transplant donors: comparison of

semiautomatic and manual methods. Radiology. 2005;234(1):171-8. Epub 2004/11/27.

85. Farraher SW, Jara H, Chang KJ, Hou A, Soto JA. Liver and Spleen Volumetry with Quantitative MR Imaging and Dual-Space Clustering Segmentation. Radiology. 2005;237(1):322-8.

86. Sahin B, Ergur H. Assessment of the optimum section thickness for the estimation of liver volume using magnetic resonance images: a stereological gold standard study. European journal of radiology. 2006;57(1):96-101. Epub 2005/08/23.

87. Gloger O, Tonies KD, Liebscher V, Kugelmann B, Laqua R, Volzke H. Prior shape level set segmentation on multistep generated probability maps of MR datasets for fully automatic kidney parenchyma volumetry. IEEE Trans Med Imaging. 2012;31(2):312-25. Epub 2011/09/23.

88. Torkzad MR, Noren A, Kullberg J. Stereology: a novel technique for rapid assessment of liver volume. Insights into imaging. 2012;3(4):387-93. Epub 2012/06/15.

89. Suzuki K, Huynh HT, Liu Y, Calabrese D, Zhou K, Oto A, et al. Computerized segmentation of liver in hepatic CT and MRI by means of level-set geodesic active contouring. Conf Proc IEEE Eng Med Biol Soc. 2013;2013:2984-7. Epub 2013/10/11.

90. Huynh HT, Karademir I, Oto A, Suzuki K. Computerized liver volumetry on MRI by using 3D geodesic active contour segmentation. Ajr. 2014;202(1):152-9. Epub 2013/12/29.

91. Kass M, Witkin A, Terzopoulos D. Snakes: Active Contour Models. International Journal of Computer Vision. 1988;1(4):pp. 321-31.

92. Falcão AX, Udupa JK, Samarasekera S, Sharma AS, Hirsch BE, Lotufo RA. User-steered image segmentation paradigms: live wire and live lane. Graph Models Image Process. 1998;60(4):233-60.

93. Heckel F, Konrad O, Hahn HK, Peitgen HO. Interactive 3D Medical Image Segmentation with Energy-Minimizing Implicit Function. Computers & Graphics. 2010;35(2):275-87.

94. Lopez-Mir F, Gonzalez P, Naranjo V, Pareja E, Alcaniz M, Solaz-Minguez J, editors. A Fast Computational Method Based on {3D} Morphology and a Statistical Filter. International Work-Conference on Bioinformatics and Biomedical Engineering IWBBIO 2013; 2013; Granada, Spain.

95. Boykov YY, Jolly MP. Interactive graph-cuts for optimal boundary & region segmentation of objects in N-D images. International Conference on Computer Vision (ICCV). 2001;I:105-12.

96. Soler L, Delingette H, Malandain G, Montagnat J, Ayache N, Koehl C, et al. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. Comput Aided Surg. 2001;6(3):131-42. Epub 2001/12/18.

97. Lamecker H, Lange T, Seebass M. Segmentation of the liver using a 3D statistical shape model. ZIB-Report 2004;04-09.

98. Nealen A, Igarashi T, Sorkine O, Alexa M. Laplacian mesh optimization Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia - GRAPHITE '06. 2006.

99. Lim SJ, Jeong YY, Ho YS. Automatic liver segmentation for volume measurement in CT images. . J Vis Commun Image R. 2006;17(4):860-75.

100. Sahin B, Mazonakis M, Akan H, Kaplan S, Bek Y. Dependence of computed tomography volume measurements upon section thickness: an application to human dry skulls. Clin Anat. 2008;21(6):479-85. Epub 2008/07/16.

101. Siewert R, Schnapauff D, Denecke T, Tolxdorff T, Krefting D. Automatic Liver Segmentation in Contrast-enhanced MRI. Bildverarbeitung für die Medizin. 2010;volume 574 of CEUR Workshop Proceedings. CEUR-WS.org:pp. 405-9.

102. Gloger O, Kuhn J, Stanski A, Volzke H, Puls R. A fully automatic three-step liver segmentation method on LDA-based probability maps for multiple contrast MR images. Magnetic resonance imaging. 2010;28(6):882-97. Epub 2010/04/23.

103. Rusko L, Bekes G. Liver segmentation for contrast-enhanced MR images using partitioned probabilistic model. International journal of computer assisted radiology and surgery. 2011;6(1):13-20. Epub 2010/06/15.

104. López-Mor F, Naranjo V, Angulo J, Alcañiz M, Luna L. Liver segmentation in MRI: A fully automatic method based on stochastic partitions. Computer methods and programs in biomedicine. 2014;114(1):11-28.

105. Wimmer A, Soza G, Hornegger J, editors. Two-stage semi-automatic organ segmentation framework using radial basis functions and level sets. . Proceedings of MICCAI 2007 Workshop: 3D Segmentation in the Clinic-A Grand Challenge,; 2007.

106. Turk G, O'Brien J. Modelling with Implicit Surfaces that Interpolate. ACM Transactions on Graphics. 2002;21(4):pp. 855-73.

107. Reiner CS, Karlo C, Petrowsky H, Marincek B, Weishaupt D, Frauenfelder T. Preoperative liver volumetry: how does the slice thickness influence the multidetector computed tomography- and magnetic resonance-liver volume measurements? Journal of computer assisted tomography. 2009;33(3):390-7.

108. Rofsky NM, Lee VS, Laub G, Pollack MA, Krinsky GA, Thomasson D, et al. Abdominal MR imaging with a volumetric interpolated breath-hold examination. Radiology. 1999;212(3):876-84. Epub 1999/09/09.

109. Chartrand G, Cresson T, Chav R, Gotra A, Tang A, de Guise JA. Liver Segmentation on CT and MR Using Laplacian Mesh Optimization. 2015.

110. Chartrand G, Chav R, Cresson T, Chantrel S, de Guise JA, editors. Live minimal path for interactive segmentation of medical images, Proc. SPIE 9413. Medical Imaging 2015: Image Processing, 94133U; 2015; Orlando, Florida.

111. Chan TF, Vese LA. Active contours without edges. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society. 2001;10(2):266-77.

112. Gotra A, Chartrand G, Massicotte-Tisluck K, Morin-Roy F, Vandenbroucke-Menu F, de Guise JA, et al. Validation of a Semiautomated Liver Segmentation Method Using CT for Accurate Volumetry. Academic radiology. 2015. Epub 2015/04/25.

113. Gotra A, Chartrand C, Vu KN, Vandenbroucke-Menu F, Massicotte-Tisluck K, de Guise JA, et al. Comparison of MRI and CT-based Semiautomated Liver Segmentation: a Validation Study 2015.

114. Foruzan AH, Chen YW, Zoroo RAF, A. , Sato Y, Hori M, Tomiyama N. Segmentation of liver in low-contrast images using K-means clustering and geodesic active contour algorithms,. IEICE Transactions on Information and Systems. 2013;E96-(D)(4):798-807.

115. Gloger O, Toennies K, Kuehn JP. Fully automatic liver volumetry using 3D level set segmentation for differentiated liver tissue types in multiple contrast MR datasets. Lecture Notes in Computer Science. Ystad, Sweden2011. p. 512-23.

116. Peng J, Wang Y, Kong D. Liver segmentation with constrained convex variational model. Pattern Recognition Letters. 2013.

117. Maklad AS, Matsuhiro M, Suzuki H, Kawata Y, Niki N, Moriyama N, et al. Blood vessel-based liver segmentation through the portal phase of a ct dataset. SPIE Medical Imaging, International Society for Optics and Photonics. 2013:86700X-X.

118. Linguraru MG, Sandberg JK, Li Z, Shah F, Summers RM. Automated segmentation and quantication of liver and spleen from CT images using normalized probabilistic atlases and enhancement estimation. Medical physics. 2010;37(2).

119. Beichel R, Bauer C, Bornik A, Sorantin E, Bischof H. Liver segmentation in CT data: A segmentation refinement approach. Proc MICCAI Workshop 3-D Segmentat Clinic: A Grand Challenge. 2007:pp. 235-45.

120. Freiman M, Eliassaf O, Taieb Y, Joskowicz L, Sosna J, editors. A bayesian approach for liver analysis: algorithm and validation study. Medical image computing and computer-assisted intervention: MICCAI; 2008:

International Conference on Medical Image Computing and Computer-Assisted Intervention 11 (Pt 1).

121. Kainm uller D, Lange T, Lamecker H, editors. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. Proc MICCAIWorkshop 3-D Segmentation Clinic: A Grand Challenge; 2007.

122. Heimann T, Munzing S, Meinzer HP, Wolf I. A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. Lecture Notes in Computer Science. Kerkrade, Netherlands2007. p. 1-12.

123. Soler L, Delingette H, Malandain G, Ayache N, Koehl C, Clement JM, et al. An automatic virtual patient reconstruction from CT-scans for hepatic surgical planning. Studies in health technology and informatics. 2000;70:316-22.

124. Aoyama M, Nakayama Y, Awai K, Inomata Y, Yamashita Y. A simple method for accurate liver volume estimation by use of curve-fitting: a pilot study. Radiological physics and technology. 2013;6(1):180-6.

125. Gotra A, Chartrand C, Vu KN, Vandenbroucke-Menu F, Massicotte-Tisluck K, de Guise JA, et al., editors. Comparison of MRI and CT-based Semiautomated Liver Segmentation: a Validation Study ARRS, American Roentgen Ray Society; 2015; Toronto, Canada.

126. Van den Broek MA, Olde Damink SW, Dejong CH, Lang H, Malago M, Jalan R, et al. Liver failure after partial hepatic resection: definition, pathophysiology, risk factors and treatment. Liver international : official journal of the International Association for the Study of the Liver. 2008;28(6):767-80. Epub 2008/07/24.

127. Pizer SM, Fletcher PT, Joshi S, Gash AG, Stough J, Thall A, et al. A method and software for segmentation of anatomic object ensembles by deformable m-reps. Medical physics. 2005;32(5):1335-45. Epub 2005/06/30.

128. Gao L, Heath DG, Kuszyk BS, Fishman EK. Automatic liver segmentation technique for three-dimensional visualization of CT data. Radiology. 1996;201(2):359-64.

129. Vauthey J-N, Abdalla EK, Doherty DA, Gertsch P, Fenstermacher MJ, Loyer EM, et al. Body surface area and body weight predict total liver volume in Western adults. Liver Transplantation. 2002;8(3):233-40.

130. Ribero D, Chun YS, Vauthey JN. Standardized liver volumetry for portal vein embolization. Seminars in interventional radiology. 2008;25(2):104-9. Epub 2008/06/01.

131. Bae KT, Giger ML, Chen CT, Kahn CE, Jr. Automatic segmentation of liver structure in CT images. Medical physics. 1993;20(1):71-8.

132. Chav R, Cresson T, Kauffmann C, de Guise JA. Method for fast and accurate segmentation processing from prior shape: application to femoral head segmentation on x-ray images. Proc of SPIE. 2009;7259:72594Y-Y-8.

133. Praun E, Hoppe H. Spherical parametrization and remeshing. ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2003. 2003;22(3):pp 340-9.

134. JM B, DG A. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;327(8476):307-10.

135. Karlo C, Reiner CS, Stolzmann P, Breitenstein S, Marincek B, Weishaupt D, et al. CT- and MRI-based volumetry of resected liver specimen: comparison to intraoperative volume and weight measurements and calculation of conversion factors. European journal of radiology. 2010;75(1):e107-11.

136. Dawant BM, Li R, Lennon B, Li S. Semi-automatic segmentation of the liver and its evaluation on the MICCAI 2007 grand challenge data set. . Proc MICCAI Workshop on 3-D Segmentat Clinic: A Grand Challenge. 2007:215-21.

137. Bauer C, Pock T, Sorantin E, Bischof H, Beichel R. Segmentation of i nterwoven 3d tubular tree structures utilizing shape priors and graph cuts. Med Image Anal. 2010 Apr;14(2):pp. 172-84.

138. Heimann T, Wolf I, Meinzer HP. Active shape models for a fully automated 3D segmentation of the liver--an evaluation on clinical data. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv. 2006;9(Pt 2):41-8. Epub 2007/03/16.

139. Lee J, Kim N, Lee H, Seo JB, Won HJ, Shin YM, et al. Efficient liver s egmentation using a level-set method with optimal detection of the initial l iver boundary from level-set speed images. Computer methods and programs in biomedicine. 2007;88(1):26-38.

140. Arya S, Mount D, Netanyahu N, Silverman R, Wu A. An optimal algorithm for approximate nearest neighbor searching. J ACM. 1998;45(6):pp. 891-923.

141. Huttenlocher D, Klanderman D, Rucklige A. Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intell. Sep. 1993;15(9):pp.850-63.

Appendix

Appendix 1. Segmentation Performance measures.

Volumetric overlap error

The volumetric overlap error (VOE) is determined using the ratio of intersection and union between two segmentations, A (automated segmentation) and M (manual segmentation) (27). It is calculated as:

$$VOE\ (A, M) = 1 - \frac{|A \cap M|}{|A \cup M|} \times 100\%$$

The VOE is 0% for a perfect overlap between segmentations and 100% for segmentations with no overlap.

Relative volume difference

The relative volume difference (RVD) between the two segmentations A and M is calculated as (27):

$$RVD\ (A, M) = \left(\frac{|A| - |M|}{|M|}\right) \times 100\%$$

A value of 0mm implies that the volumes of the two segmentations A and M are identical.

Average symmetric surface distance

The average symmetric surface distance (ASD) of surface voxels from segmentations A and M is given in millimeters. For each surface voxel from segmentation A, the Euclidean distance to the closest surface voxel of M can be

calculated (27, 140). The average of all calculated distances from A to M and M to A gives the ASD, with a perfect segmentation giving a value of 0 mm (27).

Assuming that S(A) = the set of surface voxels for semiautomated segmentation A, S(M) = set of surface voxels for manual segmentation M and the shortest distance between v (arbitrary voxel) to S (A) is:

$$d(v, S(A)) = min_{sA \in S(A)} \parallel v - s_A \parallel$$

where (‖·‖) denotes the Euclidean distance then the ASD is calculated as (27):

$$\text{ASD (A, M)} = \frac{1}{|S(A)| - |S(M)|} \left(\sum_{s_A \in S(A)} d(s_A, S(M)) + \sum_{s_M \in S(M)} d(s_M, S(A)) \right)$$

Root mean square symmetric surface distance

The root mean square symmetric surface distance (RMSD) uses the ASD previously described, however the Euclidean distances between surface voxels of A and M are squared. A perfect segmentation gives a value of 0 mm. The RMS symmetric surface distance is calculated as (27):

$$RMS(A, M) = \sqrt{\frac{1}{|S(A)| - |S(M)|}} \times \sqrt{\left(\sum_{s_A \in S(A)} d^2(s_A, S(M)) + \sum_{s_M \in S(M)} d^2(s_M, S(A)) \right)}$$

Maximum symmetric surface distance

The maximum symmetric surface distance (MSD) utilizes the maximum Euclidean distance between surface voxels from segmentations A and M (27, 141). A perfect segmentation yields a distance of 0 mm. The MSD is given in millimeters and calculated as (27):

$$MSD(A, M) = \max\left\{ max_{s_A \in S(A)} d(s_A, S(M)), max_{s_M \in S(M)} d(s_M, S(A)) \right\}$$

Appendix 2. Manuscript: Validation of a Semiautomated Liver Segmentation Method Using CT for Accurate Volumetry.

Published in *Academic Radiology* (112). Full text reproduced with permission from licensed content publisher (Elsevier) and all co-authors.

Validation of a Semiautomated Liver Segmentation Method Using CT for Accurate Volumetry

Akshat Gotra, MD, Gabriel Chartrand, B.Eng, Karine Massicotte-Tisluck, Florence Morin-Roy, MD, Franck Vandenbroucke-Menu, MD, Jacques A. de Guise, Ph.D Eng, An Tang, MD, MSc

Rationale and Objectives: To compare the repeatability and agreement of a semiautomated liver segmentation method with manual segmentation for assessment of total liver volume on CT (computed tomography).

Materials and Methods: This retrospective, institutional review board–approved study was conducted in 41 subjects who underwent liver CT for preoperative planning. The major pathologies encountered were colorectal cancer metastases, benign liver lesions and hepatocellular carcinoma. This semiautomated segmentation method is based on variational interpolation and 3D minimal path–surface segmentation. Total and subsegmental liver volumes were segmented from contrast-enhanced CT images in venous phase. Two image analysts independently performed semiautomated segmentations and two other image analysts performed manual segmentations. Repeatability and agreement of both methods were evaluated with intraclass correlation coefficients (ICC) and Bland–Altman analysis. Interaction time was recorded for both methods.

Results: Bland–Altman analysis revealed an intrareader agreement of -1 ± 27 mL (mean $\pm$ 1.96 standard deviation) with ICC of 0.999 ($P < .001$) for manual segmentation and 12 ± 97 mL with ICC of 0.991 ($P < .001$) for semiautomated segmentation. Bland–Altman analysis revealed an interreader agreement of -4 ± 22 mL with ICC of 0.999 ($P < .001$) for manual segmentation and 5 ± 98 mL with ICC of 0.991 ($P < .001$) for semiautomated segmentation. Intermethod agreement was found to be 3 ± 120 mL with ICC of 0.988 ($P < .001$). Mean interaction time was 34.3 ± 16.7 minutes for the manual method and 8.0 ± 1.2 minutes for the semiautomated method ($P < .001$).

Conclusions: A semiautomated segmentation method can substantially shorten interaction time while preserving a high repeatability and agreement with manual segmentation.

Key Words: Segmentation; liver volumetry; semiautomated; agreement; repeatability.

Acad Radiol 2015; 22:1088–1098

From the Department of Radiology, Saint-Luc Hospital, University of Montreal, 1058 rue Saint-Denis, Montreal, Quebec, Canada H2X 3J4 (A.G., K.M.-T., F.M.-R., A.T.); Department of Radiology, Montreal General Hospital, McGill University, Montreal, Quebec, Canada (A.G.); Imaging and Orthopaedics Research Laboratory (LIO), École de technologie supérieure, Centre de recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Quebec, Canada (G.C., J.A.D.G.); Department of Hepato-biliary and Pancreatic Surgery, Saint-Luc Hospital, University of Montreal, Montreal, Quebec, Canada (F.V.-M.); and Centre de recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Quebec, Canada (A.T.). Received September 23, 2014; accepted March 10, 2015. Conflicts of Interest: A.T. received a Seed Grant from the Quebec Bio-Imaging Network (QBIN #8436-0501), a New Researcher Startup grant from the Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), an operating grant from the Canadian Institute of Health Research Institute of Nutrition, Metabolism and Diabetes (CIHR-INMD #273738), and a Career Award from the Fonds de recherche du Québec en Santé (FRQS-ARQ #26993). G.C. received a MITACS industrial research award (IT02111). J.A.D.G. received funding from the Research Chair of Canada in 3D Imaging and Biomedical engineering. The other authors have no disclosures of possible conflicts of interest and/or commercial involvement. **Address correspondence to:** A.T. e-mail: [redacted]

http://dx.doi.org/10.1016/j.acra.2015.03.010

Assessment of liver volume is a mandatory step before extended hepatectomy for determining the anticipated future liver remnant and before living-donor liver transplantation for selection of appropriate candidates (1–3). Liver volumetry requires a multiplanar imaging modality. Computed tomography (CT) is currently the preferred imaging modality for surgical planning because of its superior spatial resolution and short acquisition time (4–6). Use of CT in presurgical imaging allows for concomitant assessment of vascular anatomy and quality of liver parenchyma and allows determination of total and lobar volume (7).

The reference standard method to estimate liver volume involves manually delineating the liver outline, a process called "segmentation," on consecutive CT images. This method is cumbersome, time-consuming, and impractical for widespread clinical use (8–10). Formula-based liver volume estimation using patient height and weight has also been proposed (11). However, this approach is based on a linear regression equation and is not specific to patient anatomy (12).

Automated segmentation algorithms provide several advantages such as shorter processing time, greater agreement, and repeatability (6,13–15). Although numerous studies have proposed semiautomated or fully automated liver segmentation methods from CT data sets, these methods have not necessarily been translated to clinical use (16). Reasons limiting the performance of segmentation algorithms have included small sample sizes, data sets not reflective of clinical problems, and poorly defined performance metrics (16,17). Recently, segmentation evaluation frameworks have been criticized for using liver volume alone to evaluate the quality of segmentation results (16). To facilitate the comparison between segmentation methods and objectively assess technical improvements from different research groups, several error measures have been proposed by the liver segmentation community to highlight different aspects of segmentation agreement: volumetric overlap error, average symmetric surface distance, root mean square (RMS) symmetric surface distance, and maximum symmetric surface distance (16).

In this article, we introduce a novel semiautomated liver segmentation method for CT based on variational interpolation and minimal path surface segmentation. We hypothesized that this method would improve repeatability and agreement with manual segmentation while providing faster (ie, more efficient) segmentation time. Our method is an improvement to previously published methods as no statistical shape model was imposed, which permits more segmentation flexibility for pathological or livers with unusual shape. This method is compatible with both CT and magnetic resonance (MR) data sets, which has not been previously described to our knowledge. Finally, the method is equipped with mesh-based correction tools which allow the user to achieve greater precision during interactive segmentation.

The primary aim of this study was to compare the repeatability, agreement, and efficiency of a semiautomated liver segmentation method using manual segmentation as the reference standard. The secondary aim was to evaluate the quality of segmentation using error metrics based on volume overlap and surface distances. Subsegmental volumetry was also performed based on vascular landmarks and classic anatomic principles defined by Couinaud (18).

MATERIALS AND METHODS

Our institutional review board approved this retrospective, cross-sectional study. Requirements for informed consent were waived.

Study Subjects

Our validation database consisted of 41 subjects (22 men and 19 women; mean age, 55 years) who underwent hepatectomy between October 2006 and April 2009 at our institution. Patients were included if they had primary or metastatic liver tumors and underwent major hepatectomy (three or more Couinaud segments) during the study. Each patient had a preoperative CT scan within 3 months of surgery. Hepatobiliary surgeons at our hospital independently determined indications for preoperative imaging according to clinical standard of care without influence for study inclusion. Study subjects' demographic and clinical information are summarized in Table 1.

TABLE 1. Subject Demographics

Characteristic	Data
Total subjects, *N* (%)	41 (100)
Sex	
Male (%)	22 (54)
Female (%)	19 (46)
Age (years)	
Mean ± standard deviation	55 ± 13
Body mass index in adults (kg/m^2)	
Mean ± standard deviation	26 ± 5
Pathologies	
Colorectal metastases	27 (66)
Benign liver lesions	5 (12)
Hepatocellular carcinoma	4 (10)
Cholangiocarcinoma	2 (5)
Biliary trauma	1 (2)
Cystadenocarcinoma	1 (2)
Cholangitis	1 (2)

CT Imaging Technique

CT was performed using two MDCT scanners under standard abdominal imaging protocols. Twenty-five study patients were scanned with a 16-detector scanner (Lightspeed 16; GE Medical Systems, Waukesha, WI), and 16 patients were scanned with a 64-detector scanner (Brilliance 64; Philips Medical Systems, Cleveland, OH). The parameters for the 16-detector scanner were rotation time, 0.8 seconds; detector collimation, 16 × 1.25 mm; helical pitch, 1.375; tube voltage, 120–140 kV; x-ray tube current, 75–440 mA; tube current–time product, 250 mAs. The parameters for the 64-detector scanner were rotation time, 0.75 seconds; detector collimation, 64 × 0.625 mm; helical pitch, 0.891; tube voltage, 120 kV; x-ray tube current, 151–499 mA; tube current–time product, varied based on noise index. Image reconstruction was in a 282- to 500-mm display field of view, depending on the patient's physique. Reconstruction section thickness was 2.5 mm. Reconstructed CT slices had a matrix size of 512 × 512 pixels with pixel spacing ranging from 0.55 to 0.98 mm. Before all examinations, a weight-adjusted dose of a nonionic, low-osmolar, iodinated contrast agent (Isovue; Bracco Diagnostic Inc., Princeton, NJ) was administered intravenously with a 20-ga needle at a rate of 4 mL/second. All CT protocols

included an arterial phase and portal venous phase with delays of 40 seconds and 60 seconds, respectively.

Study Workflow

The portal venous phase from the 41 subjects was used for segmentation as it provides homogeneous enhancement of the liver parenchyma and maximizes contrast between liver and nonliver structures. Liver segmentation was performed independently by four image analysts (one radiology resident, two medical students, and one biomedical engineering PhD candidate) participating in research within the department of radiology. Before this study, the image analysts received 10 hours of training in liver anatomy and software segmentation. Furthermore, the manual segmentation results used as the reference standard were supervised by an abdominal radiologist (7 years of experience). Two image analysts performed manual segmentation, whereas the other two undertook semiautomated segmentation. This ensured adequate estimation of agreement and intraobserver and interobserver repeatability. Image analysts performed repeat segmentations in a random order 1 week later to prevent recall bias. Image analysts were blinded to the results of their first segmentation and to the results of the other readers. Interaction time was recorded for both segmentation methods.

Manual Segmentation

Axial portal venous phase CT images for each patient were saved as Digital Imaging and Communications in Medicine (DICOM) files and uploaded onto imaging postprocessing display software (SliceOmatic 4.3 Rev-11; TomoVision, Montreal, Canada). For a given axial slice, two image analysts manually outlined the liver using a cursor to contour the liver. These curves then automatically deformed to precisely delineate the liver. This process generated "active contours" which are virtual curves that can be projected within images to delineate the liver boundary based on an energy equation (19). Each axial slice required further manual deformation of the active contours to completely outline the liver. Large vessels abutting the liver periphery such as the main portal vein and inferior vena cava were excluded but not vessels surrounded by liver parenchyma. The number of pixels within each contour provided the liver area on a slice-by-slice basis. This cross-sectional area was multiplied by the slice thickness and the summation of each section volume provided the total liver volume for each patient. Volumes and masks obtained from manual segmentation were used as the reference standard.

Semiautomated Segmentation and Subsegmentation

Our semiautomated segmentation method was developed at the Imaging and Orthopaedics Research Laboratory (LIO, Montreal, QC) with collaboration from the clinical and engineering teams. The method was developed and tested using MATLAB 2012a (The MathWorks, Inc., Natick, MA) computational software. Axial portal venous phase CT images for each patient were saved as DICOM files and uploaded to the segmentation program. An overview of the user and computer tasks involved in semiautomated liver segmentation is provided in Figure 1.

Initially, a seed is positioned within the liver to define a volumetric spherical region of interest used to estimate the mean intensity and standard deviation (SD). These values are used to automatically adjust the displayed contrast level and windowing to enhance the liver boundary against adjacent tissue. A liver probability density map is then generated by applying a Gaussian transfer function.

To generate an initial shape without any prior knowledge, the liver is manually delineated on one to two slices for each orthogonal plane to globally outline the liver shape. This delineation process is assisted by a snapping algorithm based on image-warping and minimal path segmentation (20). As a result, the drawn contours dynamically snap onto the liver surface.

Variational interpolation is applied to these sparse contours to interpolate a smooth surface mesh composed of vertices and triangular faces intersecting the contours initially delineated (21,22). To simplify the segmentation problem to a narrow band along the prior shape's surface, the mesh is further converted to a quadrangular mesh through surface parameterization (23). This allows the unfolding of the prior shape and the narrow band subspace which simplifies further segmentation.

The parameterized surface is then subject to two concurrent segmentations operations. First, the user can iteratively deform the mesh in three dimensions (3D) by adjusting the contours to align with actual liver anatomy. Second, the user can prompt an automated minimal surface segmentation technique to precisely delineate the liver boundary, a 3D extension of a method described by Chav et al. (20).

The final segmented mesh is converted to a volumetric mask to exclude vessel insertion points and hepatic fissures with a local region growing tool. The cross-sectional area of each mask was multiplied by the slice thickness, and the summation of each section volume provided the total liver volume for each patient.

For subsegmentation, three vertical planes were defined by drawing lines through the left, middle, and right hepatic veins and their insertion at the inferior vena cava (IVC). The portal vein bifurcation established a horizontal plane to divide segments II/III, IVa/IVb, V/VIII, and VI/VII. A polygon was then drawn to encapsulate liver tissue between the posterior aspect of the portal bifurcation and the IVC. This polygon was propagated (using an automated tool) to other slices to define the caudate lobe (Fig 2). Whole and segmental liver volumes are reported in Table 2.

Statistical Analysis

Statistical analyses were performed with SPSS software for Windows version 21.0 (Chicago, IL). Whole and segmental liver volumes were compared using paired *t* tests. Intraclass correlation coefficients (ICCs) were used to determine

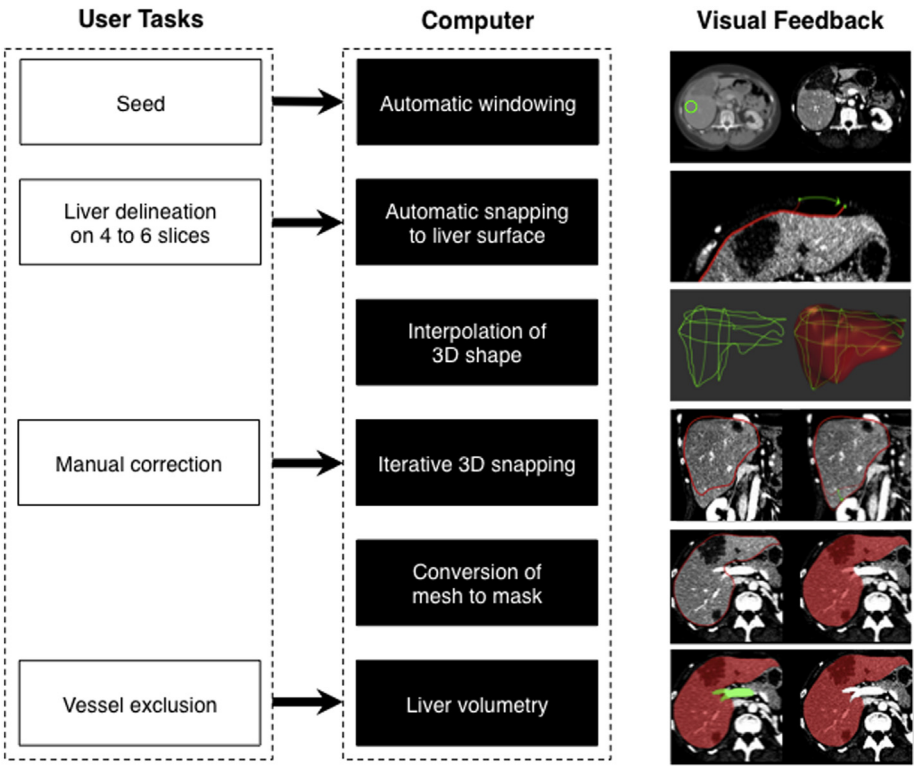

Figure 1. Overview of steps in computed tomography-based semiautomated liver segmentation. The user initiates segmentation by roughly delineating the liver contour on four to six slices. The software then uses variational interpolation to generate an initial 3D shape. This 3D shape is deformed manually then automatically by minimal path–surface segmentation. Vessels are excluded using a locally seeded region growing technique. The software then calculates liver volume for each slice. 3D, three dimensional.

intrareader, interreader, and intermethod variability of hepatic volume. Bland–Altman analyses were used to determine intrareader, interreader, and intermethod agreements. The agreement for liver volume between readers and segmentation sessions was reported as bias ± 1.96 SD of the differences, followed by the 95% limits of agreement interval (24). *P* values were calculated for Bland–Altman analyses to evaluate for systematic bias different from 0. A subgroup analysis was performed in patients with hepatocellular carcinoma (HCC). This analysis was done to determine whether the presence of underlying fibrosis or cirrhosis, which is a risk factor for HCC development, affected the results of semiautomated liver volumetry. The differences between semiautomated and manually segmented surface meshes were analyzed with four additional error measures: volumetric overlap error, average symmetric surface distance, root mean square (RMS), symmetric surface distance, and maximum symmetric surface distance (16). The formulas to calculate these segmentation error measures are reported in Appendix 1. In addition, paired *t* tests were used to compare the total interaction time for semiautomated segmentation with manual segmentation time.

RESULTS

Volumes

The mean semiautomated whole-liver volume was 1688 ± 497-mL, whereas the reference standard volume was 1689 ± 478 mL ($P = .92$). Mean segmental volumes are demonstrated in Table 2. The only statistically significant difference when comparing segmental volumetry was for the caudate lobe ($P = .01$).

Variability

Overall, eight measurements of ICC representing intrareader, interreader, and intermethod variability of hepatic volume measurements were calculated; these are summarized in Table 3. Correlation was high with an agreement between the eight ICC measures of 0.995 (95% confidence interval, 0.992–0.997). Correlation between semiautomated and manual volumetry was established with intermethod ICC values ≥0.988 ($P < .001$). Correlation for segmental volumetry readings varied greatly with values ranging from 0.331 (segment III) to 0.831 (segment VII).

Repeatability

Bland–Altman analysis showed excellent repeatability for both manual and semiautomated CT-based volumetry (Table 3). Intrareader agreement for the manual method was −6 ± 11-mL with limits of agreement of −17 and 6 mL ($P = .426$). The semiautomated method displayed higher bias of 12 ± 97 mL and wider limits of agreement of −85 and 109 mL in the repeatability calculations ($P = .291$).

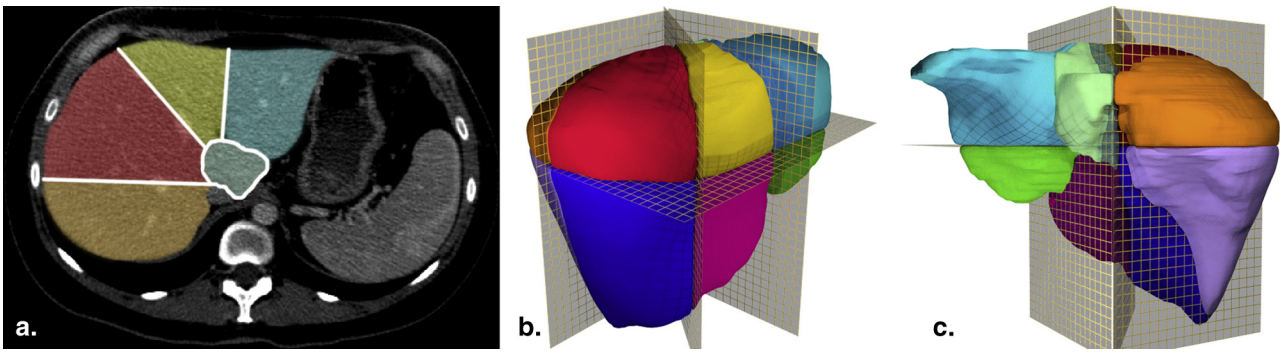

Figure 2. **(a)** Axial computed tomography slice demonstrating caudate lobe and segments II, IVa, VII, and VIII. Three vertical planes are defined by drawing lines through the left, middle, and right hepatic veins and their insertion at the inferior vena cava. A polygonal shape is propagated to define caudate lobe. **(b,c)** Oblique anterior–posterior and posterior–anterior three-dimensional renderings defining the liver subsegments.

TABLE 2. Whole and Segmental Liver Volumes by Readers

Reader	1 (Manual)	3 (Semiautomated)	*P* Value*
Whole-liver volume (mL)†	1689 ± 478	1688 ± 497	.92
Reader	**1**	**2**	***P* Value***
Segmental volume (mL)†			
I	41 ± 16	53 ± 37	.01
II	204 ± 110	186 ± 77	.31
III	97 ± 66	74 ± 57	.05
IVa	186 ± 77	205 ± 83	.18
IVb	84 ± 54	101 ± 87	.10
V	292 ± 99	278 ± 116	.26
VI	221 ± 110	202 ± 114	.14
VII	278 ± 106	292 ± 132	.24
VIII	292 ± 103	306 ± 121	.23

*Probability associated with a Student paired *t* test with a two-tailed distribution.

†Results reported as mean ± standard deviation.

Agreement

Bland–Altman analysis showed good agreement between readers for each method and between methods (Table 3). Interreader agreement for the manual method had a bias of −4 ± 22 mL with limits of agreement of −27 and 18 mL (P = .009). Interreader agreement for the semiautomated method had a bias of 5 ± 98 mL and limits of agreement of −93 and 103 mL (P = .293). Interreader agreement for segmental volumes demonstrated generally large limits of agreement, ranging from −47 and 71 (segment I) to −224 and 190 (segment II). These limits of agreement were wider than those for whole-liver volumetry.

The agreement between manual and semiautomated volumetry methods was 3 ± 120 mL with limits of agreement of −117 and 124 mL (P = .434), represented in Figure 3. Examples of concordant and discordant cases between readers are demonstrated in Figures 4 and 5, respectively.

Subgroup Analysis in Patients with HCC

Subgroup analysis in patients with HCC (n = 4) revealed correlation between semiautomated and manual volumetry with intermethod ICC values ≥0.985. Repeatability studies showed that intrareader agreement for the manual method was −4 ± 17 mL with limits of agreement of −21 and 13 mL and for the semiautomated method was 32 ± 74 mL with limits of agreement of −42 and 106 mL. The agreement between manual and semiautomated volumetry methods in this subgroup of patients was 23 ± 119 mL with limits of agreement of −96 and 142 mL.

Error Measures

Measures of segmentation agreement are summarized in Table 4. All four error calculations were slightly larger for semiautomated when compared to manual methods. Volumetric overlap error was 2.9% for manual segmentation and 4.4% for semiautomated segmentation. Overall, intermethod comparisons of manual and semiautomated segmentation yielded very low error. Volumetric overlap error was 6.4 ± 1.4% (mean ± standard deviation), average symmetric surface distance was 1.0 ± 0.2 mm, RMS symmetric surface distance was 1.8 ± 0.5 mm, and maximum symmetric surface distance was 17.0 ± 5.1 mm. Examples of 3D renderings with minimal and substantial surface distance error are shown in Figure 6.

Time

Mean interaction time was 34.3 ± 16.7 minutes per case for the manual method and 8.0 ± 1.2 minutes per case for the semiautomated method (P < .001).

DISCUSSION

This cross-sectional study evaluated the repeatability, agreement, and efficiency of a semiautomated liver segmentation

TABLE 3. Intrareader Repeatability, Interreader and Intermethod Agreement

Comparison	Readers	ICC	Bland–Altman (mL)*
Repeatability on whole-liver volume			
Intrareader	1 vs. 1	0.999	−1 ± 27 (−28, 26)
manual	2 vs. 2	1.000	−6 ± 11 (−17, 6)
Intrareader	3 vs. 3	0.995	−3 ± 67 (−70, 64)
semiautomated	4 vs. 4	0.991	12 ± 97 (−85, 109)
Agreement on whole-liver volume			
Interreader	1 vs. 2	0.999	−4 ± 22 (−27, 18)
	3 vs. 4	0.991	5 ± 98 (−93, 103)
Intermethod†	1 vs. 3	0.992	−2 ± 93 (−95, 91)
	1 vs. 4	0.988	3 ± 120 (−117, 124)
Agreement on segmental volumes			
Interreader			
Segment I	1 vs. 2	0.585	12 ± 59 (−47, 71)
Segment II	1 vs. 2	0.399	−17 ± 207 (−224, 190)
Segment III	1 vs. 2	0.331	−23 ± 139 (−162, 116)
Segment IVa	1 vs. 2	0.458	18 ± 164 (146, 182)
Segment IVb	1 vs. 2	0.713	16 ± 121 (−105, 181)
Segment V	1 vs. 2	0.758	−14 ± 150 (−164, 136)
Segment VI	1 vs. 2	0.728	−20 ± 162 (−182, 142)
Segment VII	1 vs. 2	0.831	14 ± 144 (−130, 158)
Segment VIII	1 vs. 2	0.812	14 ± 139 (−125, 153)

ICC, intraclass correlation coefficient; SD, standard deviation.

*Bland–Altman = Results reported as bias ± repeatability coefficient (1.96 SD), (95% limits of agreement interval), rounded to whole numbers.

†Intermethod agreement is reported 1 versus 3 and 1 versus 4, which represent the worst-case scenarios when comparing manual and semiautomated volumetry.

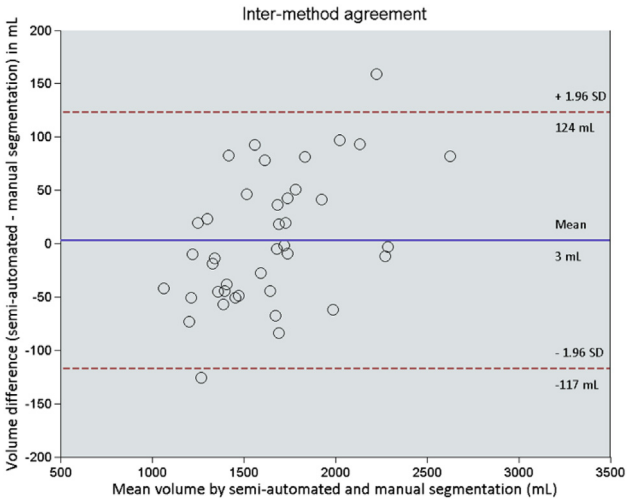

Figure 3. Bland–Altman plot of the volume difference between semiautomated and manual segmentation of computed tomography images and the mean volume (reader 1 vs. reader 4). Mean difference was demonstrated with *solid line* and 95% limits of agreement with *dashed lines*.

method by using manual segmentation as the reference standard. Overall, we found excellent correlation between semiautomated and manual segmentation volume measurements. The semiautomated method was found to have high interreader and intrareader repeatability. Furthermore, strong agreement was found between the semiautomated and manual methods when comparing volume measurements. At last, semiautomated liver volumetry was found to be time efficient.

Recently, an engineering competition comparing various liver segmentation algorithms on a common database of contrast-enhanced CT images was held (16). On average, interactive (ie, semiautomated) segmentation methods that incorporated user input were found to be more accurate and reliable than fully automated methods. The larger standard deviation of automated methods was attributed to increased outlier errors (16). The study supported the use of both statistical shape information and model-based approaches to accurately represent liver structure variability (16).

In this study, we evaluated a semiautomated segmentation method for CT images which did not require prior statistical information input. Our method represents a model-based approach and is a 3D extension of a technique developed for segmentation of femoral heads in biplane radiography (20). In introducing our novel method and validation framework, we address three limitations to segmentation performance described in the literature (17). First, we used a diverse surgical database to ensure that the method is reliable in prehepatectomy patients with a variety of hepatic pathologies. Second, we evaluated the interobserver, intraobserver, and intermethod variability in hepatic volumes. Third, we also evaluated the quality of segmentation by using volumetric and surface error measures described in the biomedical engineering literature (16). The comparison of index and reference standard segmentation meshes permits visualization of discrepancies and provides feedback for future improvement.

In our study, intrareader, interreader, and intermethod variability was assessed using ICC measures. Overall correlation was very good with an average ICC value of 0.995, indicating low variability in the measures. Semiautomated volumetry also achieved excellent correlation with manual volumetry (ICC ≥0.988). A study by Suzuki et al. (9) comparing automated and manual volumetry of living-donor livers during transplantation achieved similar results, ICC = 0.994.

Before major liver hepatectomy, the future liver remnant (FLR)-to-total liver volume ratio must be calculated (25). This ratio must be >26.5% in patients with healthy livers, >40% in patients with high-grade steatosis, and >50% in patients with cirrhosis (3). For this application, the level of agreement and reproducibility required is ±5% (26). Before living-donor liver transplantation, the liver graft-to-recipient weight ratio must be calculated. This ratio must be >0.8% and adapted to the recipient's Child class to avoid small-for-size syndrome (25). Although the level of agreement and reproducibility required for this application has not been specified, it is assumed to be the same as for major liver hepatectomy.

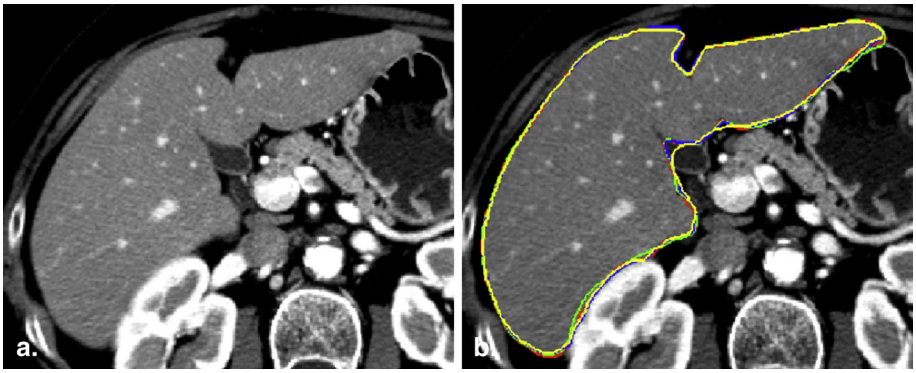

Figure 4. A 67-year-old woman with colorectal metastases. Original **(a)** and annotated **(b)**. Axial computed tomography slice demonstrating concordance between four readers using manual and semiautomated liver segmentation methods. Reader-1 manual = *red* tracing, reader-2 manual = *green* tracing, reader-3 semiautomated = *blue* tracing, reader-4 semiautomated = *yellow* tracing. (Color version of figure is available online.)

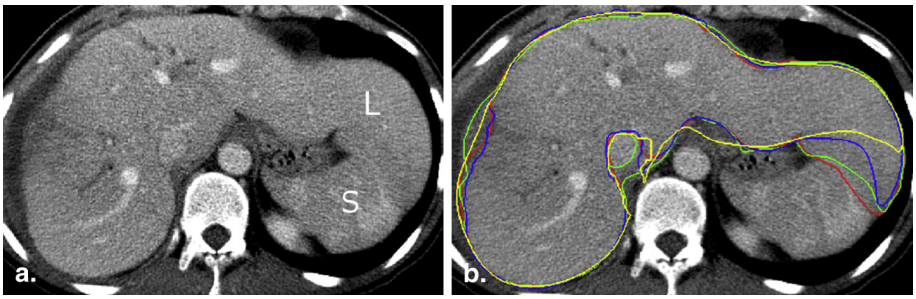

Figure 5. A 30-year-old woman with choledochal cyst. Original **(a)** and annotated **(b)**. Axial computed tomography slice demonstrating discordance between four readers using manual and semiautomated liver segmentation methods. Reader-1 manual = *red* tracing, reader-2 manual = *green* tracing, reader-3 semiautomated = *blue* tracing, reader-4 semiautomated = *yellow* tracing. Discordance between readers is found at the interface between the liver (L) and the spleen (S), the liver hilum, and the peripheral segment 8-liver lesion. (Color version of figure is available online.)

Measuring FLR volume was thought to be out of the scope of this article for a variety of reasons. Our primary aim was to accurately establish whole-liver volume as an important preliminary step before more complex segmentation procedures. Furthermore, although our data set included a variety of liver pathologies and morphologies, not all patients specifically underwent extended hepatectomy, the usual indication for calculating FLR. Finally, determining the FLR requires a clinical judgment regarding resection margin and anticipated resection plane by a hepatobiliary surgeon. For all these reasons, calculation of this parameter was not an aim of our study because it does not lead itself to automation.

Bland–Altman analysis showed excellent repeatability for both manual and semiautomated methods; however, the semiautomated method displayed higher bias. Unfortunately, we were not able to identify other studies which compared repeatability in this way for comparison. Semiautomated liver segmentation has inherent "problem regions" including the interface with adjacent structures, around blood vessels, and in the hilum of the liver which may have led to higher error. We attempted to limit such error by optimizing the initialization step of our segmentation method.

In our study, mean volume difference between readers for semiautomated segmentation was found to be 5 mL with limits of agreement of −93 and 103 mL in the Bland–Altman analysis. Similar interreader agreement was found in a study examining volumetry of resected liver specimens with achieved limits of agreement of −190 and 178 mL (27). Our narrower limits of agreement may be attributed to our readers performing segmentation on each axial slice rather than on every fourth slice as in the study by Karlo et al. (27). This potentially restricted the amount of volumetric error being interpolated to adjacent slices and led to excellent interobserver agreement in our study.

Given recent surgical advances, including increases in extended hepatectomies and split-liver and living-donor transplantation, establishing segmental and remnant liver volumetry is of growing importance. Subsegmentation was performed using classic vascular landmarks to divide the hepatic segments. Segmentation of the caudate lobe proved to be difficult as the boundaries were defined somewhat arbitrarily and not by vascular structures. Interreader correlation for segmental volumetry was found to be variable and limits of agreement were wider than those for whole-liver volumetry. This can partially be explained by our choice of portal venous phase for segmentation purposes. The hepatic veins were not always clearly visible which may have increased the subjectivity in drawing the three vertical planes. In the future, alternative acquisition phases may be acquired to facilitate subsegmentation. More reliable subsegmentation methods may also be developed based on patient-specific vascular anatomy.

TABLE 4. Segmentation Performance Measures

Error Measure	Ideal Value	Intrareader Manual (R1–R1′)*	Intrareader Semiautomated (R3–R3′)*	Intermethod (R1–R3)
Volumetric overlap error (%)	0†	2.9 ± 0.8	4.4 ± 1.3	6.4 ± 1.4
Average symmetric surface distance (mm)	0	0.4 ± 0.1	0.7 ± 0.3	1.0 ± 0.2
Root mean square symmetric surface distance (mm)	0	0.9 ± 0.2	1.6 ± 0.5	1.8 ± 0.5
Maximum symmetric surface distance (mm)	0	11.8 ± 4.9	17.2 ± 5.2	17.0 ± 5.1

Results reported as mean ± standard deviation.

*R1 and R1' indicate the first and second segmentations by reader 1, respectively. R3 and R3′ indicate the first and second segmentations by reader 3, respectively.

†0% Volumetric overlap error indicates perfect overlap between segmentation masks, whereas 100% volumetric overlap error indicates no overlap between segmentation masks.

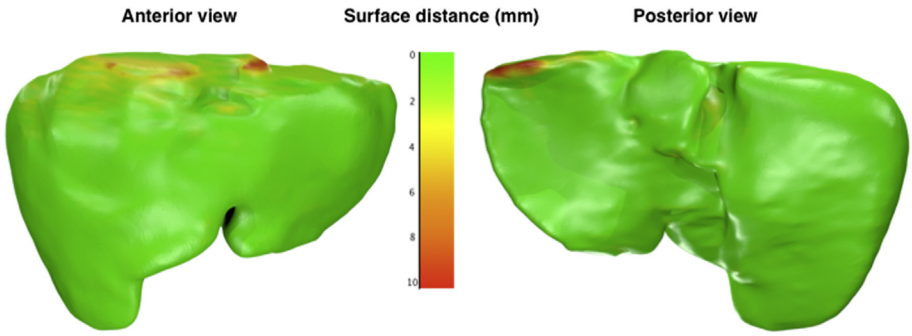

Figure 6. A 68-year-old woman with colorectal metastases. AP and PA three-dimensional renderings comparing surface distance error between semiautomated and manual segmentations. Areas in *green* represent absence of error (perfect overlap between segmentations) and areas in *red* represent surface distance error (in mm). Small amounts of error are observed at the liver dome and along the inferior vena cava. AP, anterior-posterior; PA, posterior-anterior. (Color version of figure is available online.)

Mean volume difference between semiautomated and manual CT-based volumetry methods was 3 mL with limits of agreement of −117 and 124 mL. These results are an improvement compared to recently published studies which achieved limits of agreement of −230.3 and 327 mL (5), −211 and 278 mL (9), and −503 and 509 mL (2). Our narrow limits of agreement may be attributed to a variety of factors. The small degree of user feedback during manual correction of the segmentation masks likely improved the precision of semiautomated volumetry. Furthermore, our CT-based volumetry was performed on a population of prehepatectomy patients rather than pretransplant patients. The smoother liver contours in this population, as compared to cirrhotic patients in the transplant group, may have improved agreement and precision of our method. We anticipate that a study using the same methodology as ours on a cirrhotic population may yield less-favorable results because of the more difficult segmentation inherent to nodular and dysmorphic end-stage livers. Similarly, other factors affecting hepatic parenchyma and contour, such as heterogeneous tumors, postprocedural changes, or diffuse hepatic processes may also affect volume agreement. Nakayama et al. (5) previously demonstrated that automated segmentation of damaged and deformed livers led to larger relative errors than in healthy livers. In our study, four patients had hepatocellular carcinoma. A review of their records revealed Child-Pugh scores between 5 and 7 (ie, class A or B). Furthermore, imaging did not reveal dysmorphic livers except one patient who had marked segment IV atrophy. Our study was not powered to draw comparison between cirrhotic and noncirrhotic patients by inferential statistics. Future studies on repeatability and agreement of liver segmentation may target patients with liver fibrosis or cirrhosis.

We chose to report our results according to the Bland–Altman method (23) after diligent consultation with the statistical team. Application of the Bland–Altman method for comparison between two techniques (eg, semiautomated vs. manual segmentation) is commonly used to assess "accuracy," whereas comparison of repeated measurements (eg, reader 1 vs. reader 1) is commonly used to assess "precision." The Bland–Altman method assumes (for the limits to be valid) that the error variance is constant whether expressed as a percentage or absolute value. In our article, we chose to express the error as mean differences with accompanying limits of agreement for consistency with prior literature.

Overall, use of semiautomated segmentation greatly reduced the average time required for hepatic volume determination. Mean interaction time using the semiautomated method was found to be 8.0 ± 1.2 minutes per case. This is similar to recently published studies of semiautomated liver segmentation methods which found interaction times of 20 (28), 7 (29), and 4.4 ± 1.9 minutes (5). Manual segmentation is often considered to be too time-consuming for clinical purposes (17). Thus, a fourfold decrease in mean interaction time is clinically relevant. Manual corrections within our

interactive method remained the most time-consuming step. Improving the initialization process may reduce the need for manual correction except at liver borders, where low-contrast boundaries exist with adjacent organs (14). Wider limits of agreement were noted for semiautomated than for manual segmentation. This increased variability represents a trade-off because of faster segmentation.

To compensate for the lack of specificity of volume comparison, we incorporated four novel error metrics into our segmentation evaluation framework. These metrics apply concepts of volumetric overlap and surface distance and allow for a more robust assessment of segmentation performance. Volumetric overlap error (also known as Jaccard distance) measures the dissimilarity between two segmentation results and is defined as 1 minus the ratio of intersection and union between two segmentations; a volumetric overlap error of 0% indicates perfect overlap, which is a segmentation goal. The three remaining error metrics (average, RMS, and maximum symmetric surface distance) are computed from the distribution of minimal distances between each surface point of the semiautomated segmentation and surface points from its corresponding manual segmentation; a value of 0 mm represents the ideal value for these three error metrics. Using a variety of error metrics is preferred for broad segmentation quality evaluation (16). Our volumetric overlap error of $6.4 \pm 1.4\%$ was similar to those achieved in other studies: $5.2 \pm 0.9\%$ (30), $5.8 \pm 1.4\%$ (15), and $3.8 \pm 2.2\%$ (31). Overall, our method achieved very comparable error calculations to the best interactive segmentation methods at the MICCAI 2007 Grand Challenge (16,30). At present time, there are no required performance specifications for error metrics but only ideal values which are not attainable. Yet, higher values do not disqualify automated segmentation techniques as long as they are reproducible and efficient.

When visually comparing segmentation error between readers, discordance was often found at the interface between the liver and adjacent structures (stomach, diaphragm, and body muscles), around blood vessels and in the hilum of the liver. Other studies have corroborated similar problem regions for liver segmentation. Heimann et al. (16) described segmentation error at low-contrast boundaries and near tumors. Campadelli et al. (6) described oversegmentation errors near the stomach and body muscles. Masutani et al. (4) mentioned similar density of adjacent organs as a source of error. We limited such error by adjusting windowing relative to the mean liver density.

Our study had some limitations. First, manual segmentation, as a reference standard, is not perfect. However, it is widely accepted in the literature and in standard clinical practice (2,5,9,15,16). Resected surgical liver volume or weight has also been described as alternative reference standards (5,27,32,33). However, resected specimens can provide a false estimation of in vivo liver volume because of decreased hydrostatic pressure and blood loss from the ex vivo specimens (27,32). Furthermore, CT-based volumetry methods have been shown to inaccurately estimate liver volume compared to actual surgical resection volumes (27). These physiological variations are best avoided with the use of an in vivo reference standard such as manual segmentation.

Second, we did not perform a systematic study of segmentation robustness by varying acquisition parameters such as slice thickness and injection delays (14). Yet, the purpose of our study was to simplify workflow and shorten segmentation time while maintaining good agreement (14). Third, we did not exclude all vessels in our segmentations. Standard practice remains to exclude major vessels but to include intrahepatic vessels in the total liver volume calculation (14).

In conclusion, our validation study suggests that a semiautomated liver segmentation method may provide high repeatability and strong agreement compared to manual segmentation, while substantially shortening interaction time. The quality of segmentation results was confirmed by error metrics based on overlap and surface distances. Future directions include automation of segmental volumetry based on vascular anatomy (34) and adaptation of this method to MR-based liver volumetry (35).

REFERENCES

1. Van den Broek MA, Olde Damink SW, Dejong CH, et al. Liver failure after partial hepatic resection: definition, pathophysiology, risk factors and treatment. Liver international: official journal of the International Association for the Study of the Liver 2008; 28(6):767–780.
2. Luciani A, Rusko L, Baranes L, et al. Automated liver volumetry in orthotopic liver transplantation using multiphase acquisitions on MDCT. AJR Am J Roentgenol 2012; 198(6):W568–W574.
3. Lim MC, Tan CH, Cai J, et al. CT volumetry of the liver: where does it stand in clinical practice? Clin Radiol 2014; 69(9):887–895.
4. Masutani Y, Uozumi K, Akahane M, et al. Liver CT image processing: a short introduction of the technical elements. European journal of radiology 2006; 58(2):246–251.
5. Nakayama Y, Li Q, Katsuragawa S, et al. Automated hepatic volumetry for living related liver transplantation at multisection CT. Radiology 2006; 240(3):743–748.
6. Campadelli P, Casiraghi E, Esposito A. Liver segmentation from computed tomography scans: a survey and a new algorithm. Artif Intell Med 2009; 45(2-3):185–196.
7. Kamel IR, Kruskal JB, Pomfret EA, et al. Impact of multidetector CT on donor selection and surgical planning before living adult right lobe liver transplantation. AJR Am J Roentgenol 2001; 176(1):193–200.
8. Pizer SM, Fletcher PT, Joshi S, et al. A method and software for segmentation of anatomic object ensembles by deformable m-reps. Medical physics 2005; 32(5):1335–1345.
9. Suzuki K, Epstein ML, Kohlbrenner R, et al. Quantitative radiology: automated CT liver volumetry compared with interactive volumetry and manual volumetry. AJR Am J Roentgenol 2011; 197(4):W706–W712.
10. Gao L, Heath DG, Kuszyk BS, et al. Automatic liver segmentation technique for three-dimensional visualization of CT data. Radiology 1996; 201(2):359–364.
11. Vauthey J-N, Abdalla EK, Doherty DA, et al. Body surface area and body weight predict total liver volume in Western adults. Liver Transplantation 2002; 8(3):233–240.
12. Ribero D, Chun YS, Vauthey JN. Standardized liver volumetry for portal vein embolization. Seminars in interventional radiology 2008; 25(2): 104–109.
13. Bae KT, Giger ML, Chen CT, et al. Automatic segmentation of liver structure in CT images. Medical physics 1993; 20(1):71–78.
14. Hermoye L, Laamari-Azjal I, Cao Z, et al. Liver segmentation in living liver transplant donors: comparison of semiautomatic and manual methods. Radiology 2005; 234(1):171–178.
15. Massoptier L, Casciaro S. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. European radiology 2008; 18(8):1658–1665.

16. Heimann T, van Ginneken B, Styner MA, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans Med Imaging 2009; 28(8):1251–1265.
17. Udupa JK, Leblanc VR, Zhuge Y, et al. A framework for evaluating image segmentation algorithms. Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society 2006; 30(2):75–87.
18. Couinaud C. Le Foie: études anatomiques et chirurgicales. Paris: Masson, 1957.
19. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. International Journal of Computer Vision 1988; 1(4):321–331.
20. Chav R, Cresson T, Kauffmann C, et al. Method for fast and accurate segmentation processing from prior shape: application to femoral head segmentation on x-ray images. Proc of SPIE 2009; 7259:72594Y-Y-8.
21. Heckel F, Konrad O, Hahn HK, et al. Interactive 3D medical image segmentation with energy-minimizing implicit functions. Computers & Graphics 2010; 35(2):275–287.
22. Turk G, O'Brien J. Modelling with implicit surfaces that interpolate. ACM Transactions on Graphics 2002; 21(4):855–873.
23. Praun E, Hoppe H. Spherical parametrization and remeshing. ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2003 2003; 22(3):340–349.
24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 327(8476):307–310.
25. Ben-Haim M, Emre S, Fishbein TM, et al. Critical graft size in adult-to-adult living donor liver transplantation: impact of the recipient's disease. Liver Transpl 2001; 7(11):948–953.
26. Abdalla EK, Adam R, Bilchik AJ, et al. Improving resectability of hepatic colorectal metastases: expert consensus statement. Annals of surgical oncology 2006; 13(10):1271–1280.
27. Karlo C, Reiner CS, Stolzmann P, et al. CT- and MRI-based volumetry of resected liver specimen: comparison to intraoperative volume and weight measurements and calculation of conversion factors. European journal of radiology 2010; 75(1):e107–e111.
28. Dawant BM, Li R, Lennon B, et al. Semi-automatic segmentation of the liver and its evaluation on the MICCAI 2007 grand challenge data set. Proc MICCAI Workshop on 3-D Segmentat Clinic: a grand challenge. Berlin, Heidelberg, Germany: Springer-Verlag, 2007; 215–221.
29. Lee J, Kim N, Lee H, et al. Efficient liver segmentation using a level-set method with optimal detection of the initial liver boundary from level-set speed images. Computer methods and programs in biomedicine 2007; 88(1):26–38.
30. Beichel R, Bauer C, Bornik A, et al. Liver segmentation in CT data: a segmentation refinement approach. Proc MICCAI Workshop 3-D Segmentat Clinic: a grand challenge. Berlin, Heidelberg, Germany: Springer-Verlag, 2007; 235–245.
31. Linguraru MG, Sandberg JK, Jones EC, et al. Assessing hepatomegaly: automated volumetric analysis of the liver. Academic radiology 2012; 19(5):588–598.
32. Lemke AJ, Brinkmann MJ, Schott T, et al. Living donor right liver lobes: preoperative CT volumetric measurement for calculation of intraoperative weight and volume. Radiology 2006; 240(3):736–742.
33. Yamanaka J, Saito S, Fujimoto J. Impact of preoperative planning using virtual segmental volumetry on liver resection for hepatocellular carcinoma. World journal of surgery 2007; 31(6):1249–1255.
34. Bauer C, Pock T, Sorantin E, et al. Segmentation of interwoven 3D tubular tree structures utilizing shape priors and graph cuts. Med Image Anal 2010; 14(2):172–184.
35. Siewert R, Schnapauff D, Denecke T, et al. Automatic liver segmentation in contrast-enhanced MRI. Bildverarbeitung für die Medizin. volume 574 of CEUR workshop proceedings. CEUR-WS.org. Berlin, Heidelberg, Germany: Springer-Verlag, 2010, 405–409. http://ceur-ws.org/Vol-574/.
36. Arya S, Mount D, Netanyahu N, et al. An optimal algorithm for approximate nearest neighbour searching. J ACM 1998; 45(6):891–923.
37. Huttenlocher D, Klanderman D, Rucklige A. Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intell 1993; 15(9): 850–863.

APPENDIX 1. SEGMENTATION PERFORMANCE MEASURES

Volumetric Overlap Error

The volumetric overlap error (VOE) is determined using the ratio of intersection and union between two segmentations, A (semiautomated segmentation) and M (manual segmentation). It is calculated as (16),

$$\mathrm{VOE}\ (A, M) = 1 - \frac{|A \cap M|}{|A \cup M|} \times 100\%$$

The VOE is 0% for a perfect overlap between segmentations and 100% for segmentations with no overlap.

Average Symmetric Surface Distance (ASD)

The ASD of surface voxels from segmentations A and M is given in millimeters. For each surface voxel from segmentation A, the Euclidean distance to the closest surface voxel of M can be calculated (16,36). The average of all calculated distances from A to M and M to A gives the ASD, with a perfect segmentation giving a value of 0 mm (16).

Assuming that $S(A)$ = the set of surface voxels for semiautomated segmentation A, $S(M)$ = set of surface voxels for manual segmentation M, and the shortest distance between v (arbitrary voxel) to $S(A)$ is $d(v, S(A)) = \min_{sA \in S(A)} \|v - s_A\|$, where $(\| \bullet \|)$ denotes the Euclidean distance then the ASD is calculated as (16)

$$\mathrm{ASD}(A, M) = \frac{1}{|S(A)| - |S(M)|} \left(\sum_{s_A \in S(A)} d(s_A, S(M)) + \sum_{s_M \in S(M)} d(s_M, S(A)) \right)$$

Root mean square (RMS) Symmetric Surface Distance

The RMS symmetric surface distance uses the ASD previously described; however, the Euclidean distances between surface voxels of A and M are squared. A perfect segmentation gives a value of 0 mm. The RMS symmetric surface distance is calculated as (16),

$$\mathrm{RMS}(A, M) = \sqrt{\frac{1}{|S(A)| - |S(M)|}} \times \sqrt{\left(\sum_{s_A \in S(A)} d^2(s_A, S(M)) + \sum_{s_M \in S(M)} d^2(s_M, S(A)) \right)}$$

Maximum Symmetric Surface Distance (MSD)

The MSD uses the maximum Euclidean distance between surface voxels from segmentations A and M (16,37). A perfect segmentation yields a distance of 0 mm. The MSD is given in millimeters and calculated as (16),

$$\mathrm{MSD}(A, M) = \max\left\{ \max_{s_A \in S(A)} d(s_A, S(M)), \max_{s_M \in S(M)} d(s_M, S(A)) \right\}$$