NEW METHODS TO DETECT LIVER FIBROSIS SEVERITY AND PREDICT CLINICAL OUTCOMES IN CHRONIC HEPATITIS C INFECTION

Yi Huang MBBS

This thesis is presented for the degree of Doctor of Philosophy of The University of Western Australia
School of Medicine and Pharmacology
2015
ABSTRACT

Chronic hepatitis C (CHC) is a slowly progressive disease, with an early long term asymptomatic phase. To accurately predict fibrosis severity and the risk of adverse clinical outcomes is a major challenge for individual management of CHC patients. While liver biopsy remains as the reference standard, a large number of non-invasive methods emerged as surrogate markers for liver fibrosis assessment and clinical outcome prediction. This thesis developed and validated the optimum method of image analysis to measure collagen proportional area (CPA) of liver biopsy. Novel serum models have also been developed to predict CPA and also to directly predict liver related clinical outcomes.

880 CHC patients who had a liver biopsy done in Sir Charles Gairdner Hospital from 1992 to 2012 were included in the thesis. Detailed follow up information was obtained using Western Australia data linkage unit for each patient. Clinical data including serum test results, age and gender were obtained from hospital database. The optimum method of CPA measurement was determined. Using the optimised method, CPA stage [C1: 0%-5% (normal), C2: 5%-10% (minimal), C3: 10%-20% (moderate), C4: >20% (severe)] was able to stratify risk better than Metavir stage with a significant difference in HCC free survival between each consecutive CPA stage. CPA stage remained as an independent predictor for liver related death and HCC after adjusting for Metavir stage and age. CPA stage was also significantly associated with the risk of liver related death in cirrhotic patients. Additionally, CPA measurement has the potential to use small biopsies that were insufficient for histopathological staging. A serum model
(CPAscore) was developed to detect liver fibrosis severity using CPA as the reference standard. CPAscore included HA, α2-macroglobulin and platelet count and it was closely correlated with actual CPA values with the model fit (R square) of 0.46. CPAscore achieved satisfying accuracy to detect those patients with CPA larger than 10% and 20% with AUROC of 0.82 and 0.94 respectively. Using cut points of 8.7 and 10.7, those patients with CPA larger than 10% and 20% were detected respectively with both high sensitivity and specificity. Three serum models [Liver Outcome Score (LOS)] were developed to directly predict liver related death, hepatocellular carcinoma (HCC) and liver decompensation respectively. LOS panels showed a high accuracy to predict five year liver related death, decompensation and HCC with an AUROC of 0.95, 0.90 and 0.95 respectively. Using the defined cut points, those patients categorised in the high risk group for liver related death, HCC development and decompensation had an annual incidence rate of 12.6%, 6.27% and 4.54% respectively and was significantly higher than that of low or moderate risk group (p<0.001).

In conclusion, this project it is the largest study that comprehensively evaluated the ability of CPA to predict clinical outcomes. Moreover, four serum models were developed to predict CPA value, the risk of liver related death, HCC and liver decompensation respectively. After validation, these serum models can be used as standard tests in clinical practice and guide individual patient management of chronic hepatitis C infection.
TABLE OF CONTENTS

TABLE OF CONTENTS---------------------------------------------------------------1

LIST OF FIGURES--------------------------------------------------------------------------------------------7

LIST OF TABLES-----------------------------------------------------------------------------------------------9

ABBREVIATIONS-----------------------------------------------------------------------------------------------11

ACKNOWLEDGMENTS--------------------------------------------------------------------------------------------13

CONTRIBUTION TO THESIS--------------------------------------------------------------------------------------------15

CHAPTER 1

GENERAL INTRODUCTION--------------------------------------------------------------------------------------------17

1.1. BRIEF INTRODUCTION -----------------------------------------------17

1.2. PUBLICATIONS----------------------------------------------------------19

1.2.1. Published papers--------------------------------------------------19

1.2.2. Submitted papers--------------------------------------------------19

1.2.3. Published abstracts-----------------------------------------------20

1.3. CONFERENCE PRESENTATIONS------------------------------------------23

CHAPTER 2

LITERATURE REVIEW---------------------------------------------------------------25

2.1. OVERVIEW OF CHRONIC HEPATITIS C---------------------------------------25

2.1.1. Disease burden of chronic hepatitis C------------------------------25

2.1.2. Nature history of hepatitis C virus infection-----------------------26

2.2. LIVER FIBROSIS IN CHRONIC HEPATITIS C-------------------------------29

2.2.1. Pathogenesis of liver fibrosis-------------------------------------29
2.2.2. Fibrosis progression and regression-----------------------------30
2.2.3. Cirrhosis and its complications-------------------------------31

2.3. FIBROSIS ASSESSMENT METHODS IN CHRONIC HEPATITIS C---------------------34

2.3.1. Liver biopsy and histopathological staging system-----------------35
2.3.2. Digital image analysis of liver biopsy-----------------------------36
2.3.3. Non-invasive assessment of liver fibrosis--------------------------39
2.3.4. Serum tests----------------------------------------------------------39
  2.3.4.1. HA---------------------------------------------------------------40
  2.3.4.2. PIIINP and type IV collagen--------------------------------------41
  2.3.4.3. MMPs and TIMPS-----------------------------------------------40
  2.3.4.4. Serum models-----------------------------------------------------42
2.3.5. Radiologic tests-----------------------------------------------------47
  2.3.5.1. Ultrasound-based transient elastography------------------------47
  2.3.5.2. Acoustic radiation force impulse imaging------------------------48
  2.3.5.3. Real-time shear wave elastography-------------------------------49
  2.3.5.4. Magnetic resonance elastography--------------------------------50
  2.3.5.5. Diffusion-weighted magnetic resonance imaging------------------51
2.3.6. Combining tests-----------------------------------------------------52

2.4. OUTCOME PREDICTION FOR CHRONIC HEPATITIS C------------------------53

2.4.1. Outcome predictors for cirrhotic patients--------------------------53
  2.4.1.1. Clinical manifestations-------------------------------------------54
  2.4.1.2. Extra histopathological features and image analysis---------------55
  2.4.1.3. Hepatic venous pressure gradients-------------------------------56
  2.4.1.4. Serum markers and models----------------------------------------57
  2.4.1.5. Ultrasound-based transient elastography-------------------------60
2.4.2. Outcome prediction beyond cirrhotic patients-----------------------61
  2.4.2.1. Histopathological fibrosis stage-------------------------------62
  2.4.2.2. Serum markers---------------------------------------------------62
  2.4.2.3. Transient elastography-------------------------------------------64

2.5. SUMMARY STATEMENT----------------------------------------------------66

2.6. AIMS OF THE THESIS---------------------------------------------------69
CHAPTER 3

CLINICAL OUTCOMES OF CHRONIC HEPATITIS C PATIENTS RELATED TO BASELINE LIVER FIBROSIS STAGE – A HOSPITAL BASED LINKAGE STUDY

3.1. ABSTRACT

3.1.1. Background and aims

3.1.2. Methods

3.1.3. Results

3.1.4. Conclusions

3.2. INTRODUCTION

3.3. METHODS

3.3.1. Patients

3.3.2. Liver fibrosis evaluation

3.3.3. Outcome measurement

3.3.4. Statistical methods

3.4. RESULTS

3.5. DISCUSSION

CHAPTER 4

IMAGE ANALYSIS OF LIVER COLLAGEN USING SIRIUS RED IS MORE ACCURATE AND CORRELATES BETTER WITH SERUM FIBROSIS MARKERS THAN TRICHROME

4.1. ABSTRACT

4.1.1. Background

4.1.2. Aims

4.1.3. Methods

4.1.4. Results

4.1.5. Conclusions

4.2. INTRODUCTION

4.3. METHODS

4.3.1. Patients
CHAPTER 8

GENERAL DISCUSSION---------------------------------------------------------------169

8.1. SUMMARY AND DISCUSSION-----------------------------------------------------169
8.2. FUTURE DIRECTIONS----------------------------------------------------------176
   8.2.1. Image analysis of liver biopsy------------------------------------------176
   8.2.2. CPA measurement------------------------------------------------------177
   8.2.3. Serum models-----------------------------------------------------------178
      8.2.3.1. CPAscore----------------------------------------------------------178
      8.2.3.2. Liver Outcome Score panel------------------------------------------178

REFERENCES---------------------------------------------------------------------181
LIST OF FIGURES

FIGURE 3.1. FLOW CHART OF PATIENT INCLUSION AND FOLLOW UP---81

FIGURE 3.2. LIVER RELATED SURVIVAL FOR COMPENSATED PATIENTS WITH EACH METAVIR STAGE-----------------------------------------------82

FIGURE 3.3. CLINICAL OUTCOMES FOR COMPENSATED PATIENTS WITH EACH METAVIR STAGE---------------------------------------------------------------------------------------------83

FIGURE 3.4. LIVER RELATED SURVIVAL FOR PATIENTS WITH LIVER COMPLICATION-----------------------------------------------------------------------------------------------83

FIGURE 4.1. CPAS VARIATION ASSOCIATED WITH SIZE OF MEASUREMENT AREA USING BIOPSIES FROM TWO PATIENTS----------102

FIGURE 4.2. CPAT AND CPAS FOR EACH PATIENT. PATIENTS WERE SORTED ACCORDING TO CPAS--------------------------------------------------------------------------------------------103

FIGURE 4.3. LIVER BIOPSIES STAINED USING SIRIUS RED AND TRICHRome FOR EACH METAVIR STAGE----------------------------------------------------------------------------------104

FIGURE 4.4. CPAS AND CPAT GROUPED ACCORDING TO METAVIR STAGE IN 201 CHC PATIENTS-----------------------------------------------------------------------------------105

FIGURE 5.1. OVERLAP OF CPA VALUES AND METAVIR STAGE--------124

FIGURE 5.2. CLINICAL OUTCOMES ACCORDING TO CPA STAGE--------125

FIGURE 5.3. CLINICAL OUTCOMES ACCORDING TO METAVIR STAGE-126

FIGURE 6.1. CORRELATIONS BETWEEN CPASCORE AND ACTUAL CPA VALUE---------------------------------------------------------------------------------------------------------------142
FIGURE 6.2. ROC CURVES FOR CPASCORE-----------------------------143

FIGURE 7.1. SURVIVAL CURVES ACCORDING TO LOS PANEL----------161

FIGURE 7.2. ROC CURVES OF SERUM MODELS TO PREDICT HCC DEVELOPMENT----------------------------------162
LIST OF TABLES

TABLE 2.1. LIST OF SERUM MODELS -------------------------------------------44

TABLE 2.2. PREDICTIVE PERFORMANCE OF SERUM MODELS------------------45

TABLE 2.3. CHILD-TURCOTTE-PUGH SCORE---------------------------------58

TABLE 3.1. PATIENT CHARACTERISTICS--------------------------------------84

TABLE 3.2. HAZARD RATIO FOR LIVER COMPLICATION AND LIVER RELATED DEATH----------------------------------------------------------84

TABLE 3.3. LIVER RELATED SURVIVAL PROBABILITY FOR EACH METAVIR STAGE------------------------------------------------------------85

TABLE 3.4. HCC FREE SURVIVAL PROBABILITY FOR EACH METAVIR STAGE---------------------------------------------------------------85

TABLE 3.5. LIVER DECOMPENSATION FREE SURVIVAL PROBABILITY FOR EACH METAVIR STAGE-----------------------------------------------86

TABLE 4.1. PATIENT CHARACTERISTICS----------------------------------------106

TABLE 4.2. CORRELATION BETWEEN SERUM MARKERS AND CPA---------------------107

TABLE 4.3. MULTIVARIATE ANALYSIS OF SERUM MARKERS AND CPA---------------------------------------------------------------------------108

TABLE 4.4. CORRELATION BETWEEN SERUM MODELS AND CPA---------------------108

TABLE 5.1. PATIENTS' CHARACTERISTICS (N=533)----------------------------127
TABLE 5.2. PREDICTORS FOR DECOMPENSATION, HCC AND LIVER RELATED DEATH

TABLE 5.3. 15 YEAR SURVIVAL PROBABILITY OF EACH END POINT FOR EACH CPA STAGE

TABLE 6.1. PATIENT CHARACTERISTICS

TABLE 6.2. CORRELATIONS BETWEEN SERUM MARKERS AND CPA

TABLE 6.3. MODELS TO PREDICT CPA

TABLE 7.1. PATIENTS CHARACTERISTICS

TABLE 7.2. UNIVARIATE AND MULTIVARIATE ANALYSIS TO PREDICT EACH END POINT

TABLE 7.3. INCIDENT RATES FOR EACH ENDPOINT ACCORDING TO CORRESPONDING MODELS

TABLE 7.4. COMPARISON OF THE ACCURACY OF LOS PANEL WITH OTHER MODELS IN PREDICTING LIVER OUTCOMES
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHC</td>
<td>Chronic hepatitis C</td>
</tr>
<tr>
<td>ALD</td>
<td>Alcoholic liver disease</td>
</tr>
<tr>
<td>ALP</td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td>ALT</td>
<td>Alanine aminotransferase</td>
</tr>
<tr>
<td>APRI</td>
<td>AST to platelet ratio index</td>
</tr>
<tr>
<td>AST</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>AUROC</td>
<td>Area Under Receiver Operating Characteristic</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CLD</td>
<td>Chronic liver disease</td>
</tr>
<tr>
<td>CPA</td>
<td>Collagen proportional area</td>
</tr>
<tr>
<td>CPAs</td>
<td>CPA measured using sirius red stained liver biopsy</td>
</tr>
<tr>
<td>CPAt</td>
<td>CPA measured using trichrome stained liver biopsy</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of variance</td>
</tr>
<tr>
<td>DIA</td>
<td>Digital image analysis</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>ELF</td>
<td>Enhanced Liver Fibrosis score</td>
</tr>
<tr>
<td>GGT</td>
<td>Gamma-glutamyl transpeptidase</td>
</tr>
<tr>
<td>GUCI</td>
<td>Goteborg University Cirrhosis Index</td>
</tr>
<tr>
<td>HA</td>
<td>Hyaluronic acid</td>
</tr>
<tr>
<td>HBV</td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>HCC</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>insulin resistance by the Homeostasis model assessment</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>HSC</td>
<td>Hepatic stellate cells</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>HVPG</td>
<td>Hepatic venous pressure gradient</td>
</tr>
<tr>
<td>ICC</td>
<td>Inter-class correlation</td>
</tr>
<tr>
<td>INR</td>
<td>International normalized ratio</td>
</tr>
<tr>
<td>LOS</td>
<td>Liver Outcome Score</td>
</tr>
<tr>
<td>MELD</td>
<td>Model for End-stage Liver Disease</td>
</tr>
<tr>
<td>MMP</td>
<td>Matrix metalloproteases</td>
</tr>
<tr>
<td>PIIINP</td>
<td>Procollagen III N-peptide</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SF</td>
<td>Significant fibrosis</td>
</tr>
<tr>
<td>TIMPs</td>
<td>Tissue inhibitors of matrix metalloproteinase</td>
</tr>
<tr>
<td>VH</td>
<td>Viral hepatitis</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

Three years ago, when I just arrived at Australia, I had nothing but the courage to start my PhD and the luck to have Winthrop Professor Gary Jeffrey as my coordinate supervisor. I would like to express my sincere thanks to him for his encouragement and trust that guided me into research, and for his care and numerous efforts to offer guidance when I faced trouble. It was his wisdom, enthusiasm and kindness that deeply influenced me and made my PhD an enjoyable journey.

There are many other people to thank for the completion of this thesis. I would like to thank my supervisor Professor Gerry MacQuillan, Professor Leon Adams and Professor Max Bulsara for their helpful guidance and suggestions throughout my PhD. I would like to express my gratitude to Doctor Bastiaan de Boer for his great support in liver biopsy analysis. I also like to thank Doctor Enrico Rossi for his help in serum tests collection and Professor Paul Rigby for providing the facilities of image analysis. I am thankful to all the doctors, nurses and faculties in Department of Hepatology, Sir Charles Gardner Hospital and PathWest who helped me in each step of my project. A special thanks to my colleagues and friends, who listened again and again about my project, cheered me up when I was down and shared the happiness with me when I had an achievement.

PhD is never an individual achievement but a group work. Without any of them, I wouldn’t be able to complete my PhD thesis as smoothly as now I have.
During these three years, I have gained much more than a PhD itself, and it is the other things that I cherish the most.
## CONTRIBUTION TO THESIS

<table>
<thead>
<tr>
<th>Name</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>W/Professor Gary Jeffery</td>
<td>Coordinator supervisor, assisted with study design, data collection, interpretation and manuscript preparation</td>
</tr>
<tr>
<td>Professor Leon Adams</td>
<td>Co-supervisor, assisted with study design and data interpretation</td>
</tr>
<tr>
<td>Clin/Professor Gerry MacQuillan</td>
<td>Co-supervisor, assisted with study design and data interpretation</td>
</tr>
<tr>
<td>Professor Max Bulsara</td>
<td>Co-supervisor, assisted with data analysis and interpretation</td>
</tr>
<tr>
<td>Doctor Bastiaan de Boer</td>
<td>Assisted with histopathological staging of liver biopsy</td>
</tr>
<tr>
<td>Doctor Enrico Rossi</td>
<td>Facilitated serum tests collection</td>
</tr>
<tr>
<td>Professor Paul Rigby</td>
<td>Facilitated image analysis of liver biopsy</td>
</tr>
</tbody>
</table>
CHAPTER 1

GENERAL INTRODUCTION

1.1. BRIEF INTRODUCTION

The primary goal of this project was to develop and validate new methods to predict liver fibrosis severity as well as clinical outcomes for CHC patients. More specifically, the prognostic performance of image analysis of liver biopsy was evaluated and compared with that of the reference standard - histopathological staging system. A serum model was developed to predict liver fibrosis severity using image analysis as gold standard. Additionally, serum models were developed to directly predict clinical outcomes for CHC patients.

The main body of this work was presented in detail in Chapter 3 to 7. In chapter 3, the long term clinical outcome for CHC patients according to histopathological fibrosis stage (Metavir stage) was evaluated. In chapter 4 and 5, the optimal method for image analysis was determined and the prognostic ability of image analysis was evaluated and compared with that of Metavir stage. In chapter 6, a serum model was developed to predict liver fibrosis severity using image analysis as reference standard. In chapter 7, three serum models were developed to directly predict each adverse clinical outcome, namely: liver related death, liver decompensation and hepatocellular carcinoma.

This project added important new evidence to the existing literature. It was the largest study to date that evaluated the clinical utility of image analysis.
Furthermore, serum models were developed to predict liver fibrosis severity and clinical outcomes and these will provide important information to guide patient prognosis and management.
1.2. PUBLICATIONS

1.2.1. Published papers

Huang Y, de Boer WB, Adams LA, MacQuillan G, Bulsara KM, Jeffrey GP. Clinical outcomes of chronic hepatitis C patients related to baseline liver fibrosis stage – a hospital based linkage study. Intern Med J 2014 (accepted) (Incorporated into chapter 3)


1.2.2. Submitted papers

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Bulsara KM, Jeffrey GP. A new serum biochemical model that accurately predicts liver collagen proportional area. (Incorporated into chapter 6)

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Bulsara KM, Jeffrey GP. New simple serum marker models reliably predict liver related clinical outcomes in chronic hepatitis C infection. (Incorporated into chapter 7)
1.2.3. Published abstracts

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Bulsara KM, Jeffrey GP. Simple serum markers models accurately predict liver related survival, complications and HCC in chronic HCV infection. 2014 The Liver Meeting, AASLD, Boston, USA

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Bulsara KM, Jeffrey GP. A new serum biochemical model that accurately predicts liver collagen proportional area. 2014 The Liver Meeting, AASLD, Boston, USA

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Bulsara KM, Jeffrey GP. Simple serum markers models accurately predict liver related survival, complications and HCC in chronic HCV infection. 2014 Australia Gastroenterology Week, Gold coast, Australia

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Bulsara KM, Jeffrey GP. A new serum biochemical model that accurately predicts liver collagen proportional area. 2014 Australia Gastroenterology Week, Gold coast, Australia

Huang Y, de Boer WB, Adams LA, MacQuillan G, Bulsara KM, Jeffrey GP. Semi-automated digital scan technology allows fibrosis measurement in small liver biopsy samples that accurately correlates with clinical outcomes. 2013 Australia Gastroenterology Week, Melbourne, Australia
Huang Y, de Boer WB, Adams LA, MacQuillan G, Bulsara KM, Jeffrey GP. Clinical outcomes of chronic hepatitis C patients related to baseline liver fibrosis stage – a hospital based linkage study. 2013 Australia Gastroenterology Week, Melbourne, Australia

Huang Y, de Boer WB, Adams LA, MacQuillan G, Bulsara KM, Jeffrey GP. Semi-automated digital scan technology allows fibrosis measurement in small liver biopsy samples that accurately correlates with clinical outcomes. 2012 The Liver Meeting, AASLD, Washington, USA.

Huang Y, de Boer WB, Adams LA, MacQuillan G, Bulsara KM, Jeffrey GP. Clinical outcomes of chronic hepatitis C patients related to baseline liver fibrosis stage – a hospital based linkage study. 2012 The Liver Meeting, AASLD, Washington, USA.

Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Rigby P, Raftopoulos SC, Jeffrey GP. Collagen proportional area using Sirius red stained liver biopsies is more accurate than Masson’s trichrome staining and reliably predicts the development of hepatocellular carcinoma. 2012 The Liver Meeting, AASLD, Boston, USA

Huang Y, de Boer WB, Adams LA, MacQuillan G, Jeffrey GP. Image Analysis of Liver Collagen using Sirius Red Staining is more Accurate than Trichrome and Reliably Predicts the Development of Hepatocellular Carcinoma. 2012 Combined Biological Sciences Meeting, Perth Australia
Huang Y, de Boer WB, Adams LA, MacQuillan G, Rossi E, Rigby P, Raftopoulos SC, Jeffrey GP. Collagen proportional area using Sirius red stained liver biopsies is more accurate than Masson’s trichrome staining and reliably predicts the development of hepatocellular carcinoma. 2012 Australia Gastroenterology Week, Adelaide, Australia
1.3. CONFERENCE PRESENTATIONS

2014 Australia Gastroenterology Week, gold coast, Australia (poster)
2014 The Liver Meeting, Boston, USA (poster)
2013 Australia Gastroenterology Week, Melbourne, Australia (oral & poster)
2013 International Liver Transplantation Society, Sydney, Australia (oral)
2013 The Liver Meeting, Washington, USA (poster)
2013 West Coast Liver Meeting, Margaret River, Australia (oral)
2012 Australia Gastroenterology Week, Adelaide, Australia (oral)
2012 The Liver Meeting, Boston, USA (poster)
2012 Combined Biological Sciences Meeting, Perth Australia (poster)
CHAPTER 2

LITERATURE REVIEW

2.1. OVERVIEW OF CHRONIC HEPATITIS C

2.1.1. Disease burden of chronic hepatitis C

About 170 million individuals are chronically infected with hepatitis C virus (HCV) worldwide and three to four million people are newly infected each year (1). Chronic hepatitis C infection contributes to 27% cases of liver cirrhosis, 25% cases of HCC and 350,000 death per year (2). In Australia, HCV infection affects 300,000 people with 2550 CHC related death per year (3). Intravenous drug use with shared needles is the main source of HCV transmission, accounting for more than 60% of the cases. Blood transfusion was another main transmission source before 1992 but it has been virtually eliminated after the introduction of routine testing of donated blood (4). Other risk factors for HCV infection include tattooing, body piercing, hospitalisation, organ transplantation and sexual contact (5-7).

The prevalence of HCV infection peaked at the year 2001 and started to decrease afterwards (8). Although the decline in disease prevalence and incidence of newly HCV infection is observed in most developed countries, the proportion of cases with advanced fibrosis will continue to increase during the next two decades. The number of cases of cirrhosis is projected to peak after the year 2020 (8). Additionally, majority of HCV infected population reach the ages at which complications from chronic liver disease typically occur. In
Australia, about 70% of CHC patients have an age of 40 or more in 2012 (3). This suggests that the incidence of HCV related complications, HCC and death will continue to increase in the next two decades (9).

Chronic hepatitis C imposes a significant economic burden in health care system worldwide. The annual direct cost for hepatitis C was estimated to be $694–$1660 million in America, $140 million of which was associated with HCC and the remaining cost due to HCV treatment, liver complication management, and liver transplantation (10). In Australia, the annual expenditure for liver disease was estimated to be $386.2 million in 2012: $303 million for hospital admitted patient services, $63.2 million for out of hospital medical services and $19.9 million for prescription pharmaceuticals (3). Moreover, the indirect cost for hepatitis C, including the cost of forgone production because of hospitalization, ambulatory care, premature death, and work loss, was estimated to be 67% higher than estimated direct costs (10).

### 2.1.2. Nature history of hepatitis C virus infection

Most acute HCV infection cases are asymptomatic and undetectable. The viral infection, however, usually persists and leads to the development of chronic hepatitis. Spontaneous viral clearance rate varies from 14% to 50% among studies and this is associated with the mode of HCV acquisition, race and comorbidities such as HIV co-infection (11-14).

The major consequences of chronic hepatitis C are the development of liver decompensation or HCC and these eventually lead to the requirement for liver
transplantation or death. The disease prognosis is closely associated with the severity of liver fibrosis and the majority of adverse outcomes occur after the development of liver cirrhosis (15). The natural history of chronic hepatitis C is difficult to define due to the prolonged course of the disease and the uncertainty of infection acquisition time. It is generally accepted that chronic hepatitis C is a slowly progressing disease with a minimal risk of liver related morbidity and mortality in the first two decades after infection. Studies followed patients from infection to around 20 years found that the rate of HCC development and liver related death was 0-2.9% and 0.2-3.7% respectively (16-18). A Meta-analysis of 111 studies found that 16% people developed cirrhosis twenty years after infection (19). Disease progression beyond twenty years was less well documented. The rate of HCC development and liver related death in young German women infected with HCV following anti-D globulin use was 1.5% and 0.5% respectively after 25 years of infection (12). One study reported worse outcomes with hepatocellular carcinoma development rate of 4.7% and liver related death rate of 5.3% after 31 years of infection (20). Another study followed 17 young men for 45 years and found one (5.9%) had liver related death, two (11.8%) had liver related morbidity, and no one had cancer (21).

Chronic hepatitis C is also a heterogenetic disease with a number of well documented host and viral factors that influence disease progression (22). Evidence showed that cirrhosis development rate in 20 years was less than 12% for CHC patients infected before the age of 30 and was more than 50% for those infected after 40 years old (23). A Meta-analysis found that the mean 20 year cirrhosis prevalence was 14.8% among individuals infected with HCV through intravenous drug use and a 5% increase was expected in male patients,
those with excessive alcohol consumption and those with HIV co-infection (24). Other host factors such as race, infection acquisition source and genetic factors also potentially influence disease progression rate (25-28). The effect of viral factors on disease progression is less clear. HCV viral load appears to remain at a stable level for each individual but doesn’t significantly associate with disease progression rate (23). Contradictory results were found on the effect of viral genotype on disease progression (23, 29-31).
2.2. LIVER FIBROSIS IN CHRONIC HEPATITIS C

2.2.1. Pathogenesis of liver fibrosis

Liver fibrogenesis occurs in response to chronic liver injury and results in the change in the amount, distribution and quality of extracellular matrix in the liver. Over time, the amount of extracellular matrix increase to up to 8-10 fold over that presented in normal liver and the matrix composition changes from one comprised of type IV collagen, heparansulfate proteoglycan, and laminin to one rich in fibril-forming collagens, particularly types I and III (32). The distribution of newly formed extracellular matrix is abnormal and eventually leads to complete loss of the normal architecture of the liver.

Hepatic stellate cells (HSC) are the primary source of extracellular matrix (ECM) in normal and fibrotic liver. They locate in subendothelial space of Disse between hepatocytes and sinusoidal endothelial cells and constitute 10–15 % of the total number of resident cells in normal liver (33). Under physiological conditions, stellate cells reside in a quiescent stage and play an important role in vitamin A (retinol) transport and storage. Following liver injury, hepatic stellate cells undergo an activation process and transition from quiescent vitamin A-rich cells into highly proliferative, fibrogenic and contractile myofibroblasts. Activated hepatic stellate cells synthesize and secrete a large amount of extracellular matrix components such as collagen, proteoglycan, glycosaminoglycan, and glycoprotein (34). Apart from activated stellate cells, other cell type such as portal fibroblasts and circulating fibrocytes, also contributed to extracellular matrix deposition (35, 36). This is further enhanced by the production of tissue inhibitors of matrix metalloproteinase (TIMPs), which prevent the degradation of...
extracellular matrix, leading to a net accumulation of extracellular matrix with a gradual disruption of normal liver architecture (37).

Angiogenesis and disruption of liver vascular architecture are closely associated with progressive fibrogenesis. The overexpression of growth factors (such as platelet derived growth factor, transforming growth factor-b1, fibroblast growth factor and vascular endothelial growth factor) and the increased tissue hypoxia in fibrotic liver are the main triggers of new blood vessels formation (38). As a consequence, intrahepatic shunts develop within the liver and this leads to increased hepatic resistance and decreased effective hepatocyte perfusion which further results in portal hypertension and liver failure respectively (38).

2.2.2. Fibrosis progression and regression

Liver fibrosis in hepatitis C progresses in a slow manner with about 30% of patients won't progress to cirrhosis in their life time (39). Fibrosis progression rate has been estimated either as the ratio of fibrosis stage and the total infection time or as the ratio of fibrosis change between two liver biopsies and the interval time. One of the biggest studies included 2235 CHC patients and demonstrated a significant correlation between infection time and fibrosis stages. The rate of fibrosis progression was estimated to be 0.13 Metavir stage per year (39). Another study included 123 patients who had two liver biopsies with a mean interval time of 44 months and found 39% patients had fibrosis progression, 37% had no change, and 24% had improvement. The mean rate of fibrosis progression was 0.12 Ishak stage per year (40). However, these estimations were based on the assumption that liver fibrosis progress rate remained constant over time. A more recent study proposed a Markov model
that was able to evaluate the fibrosis progression rate for each fibrosis stage (41). Using this model, a Meta-analysis found that the fibrosis progression was not linearly correlated with infection time with a stage specific annual transition probability of 0.117 for F0-F1, 0.085 for F1-F2, 0.120 for F2-F3 and 0.116 for F3-F4 (19).

Increasing evidence showed that liver fibrosis and potentially early cirrhosis can regress especially during or after antiviral treatment. One study compared the pre-treatment and post-treatment liver biopsy for 3010 CHC patients and found that 25% patients with sustainable viral clearance had fibrosis regression (42). A rate of 49% of cirrhosis reversion for sustained responders was also observed (42). Another study including 287 CHC patients reported higher fibrosis regression rate of 41% during antiviral treatment for sustained responders (43). Other studies followed sustainable responders after treatment for three to five years and found that fibrosis regression occurred in 44% to 80% of patients (44-46).

2.2.3. Cirrhosis and its complications

Cirrhosis is the most advanced stage of liver fibrosis and is characterized by the distortion of the liver parenchyma associated with nodule formation, altered blood flow, and dramatically increased risk of liver complications and death. In clinical terms, cirrhosis is commonly categorised into compensated cirrhosis and decompensated cirrhosis. Compensated cirrhosis is the asymptomatic stage of cirrhosis. It progresses relatively slowly and most commonly transitions to decompensated cirrhosis before death. At compensated cirrhosis phase, the annual incident rate of liver decompensation, hepatocellular carcinoma and
death is estimated to be 6.37%, 3.36% and 4.58% respectively (47). Decompensated cirrhosis occurs when clinical evident liver complication develops and it rapidly progress towards death or liver transplantation.

Portal hypertension is the earliest and the most common consequence of cirrhosis. It is observed in around 80% to 90% of totally asymptomatic cirrhotic patients (48). Elevated portal pressure leads to splenomegaly, a growth of extensive network of portal-systemic collaterals and the development of portal-systemic blood flow shunts. These further contribute to the development of liver decompensation. Ascites is the most common form of liver decompensation followed by variceal bleeding (49). Other decompensation symptoms include encephalopathy, hepatorenal syndrome, spontaneous bacterial peritonitis, hepatic hydrothorax, hepatopulmonary syndrome, and non-obstructive jaundice.

The risk of HCC development is eleven times higher in CHC patients than non-CHC population (50). It is the leading cause of liver related death in CHC patients, accounting for 44% of the cases (51). HCC occurs predominantly in cirrhotic patients with rare cases that it develops before cirrhosis (52). The mechanism of HCC development in chronic hepatitis C is not entirely clear. It is generally accepted that the oxidative stress and the continuous inflammation induced by HCV facilitate the accumulation of genetic alterations within the hepatocyte and eventually lead to HCC development (53). Additionally, various HCV proteins such as core protein, envelope protein and non-structural protein are proved to disturb cellular regulatory pathways that associate with proliferation and apoptosis, and suppress host immune responses (54, 55). Risk factors such as alcohol intake, co-infection with Hepatitis B virus (HBV) or
human immunodeficiency virus (HIV) and diabetes also predispose those patients to HCC development (56, 57).
2.3. FIBROSIS ASSESSMENT METHODS IN CHRONIC HEPATITIS C

As the prognosis of chronic hepatitis C is closely associated with liver fibrosis severity, a precise estimation of the degree of liver fibrosis is of great importance for patient management. Two stages of fibrosis severity are critical in clinical practice: significant fibrosis (Metavir F2-F3 or Ishak stage 2-6) and cirrhosis (Metavir F4 or Ishak stage 5-6). The presence of significant fibrosis suggests a reasonable risk of developing CHC associated morbidities in one’s lifetime and it is when antiviral treatment is recommended in most countries. The development of cirrhosis marks the dramatically increased risk of liver decompensation, HCC development and liver related death and it is when routine surveillances of esophageal varicies and HCC start (58).

Liver biopsy has long been used as the gold standard to detect liver fibrosis severity. In addition to fibrosis assessment, liver biopsy can provide extra information such as necroinflammation, steatosis, iron accumulation and other aetiologies of chronic liver disease (59). However, because of the invasive manner of liver biopsy, it is poorly tolerated especially for multiple tests in longitudinal settings. As a result, a large number of non-invasive methods for liver fibrosis assessment have been developed as surrogate markers for liver biopsy. They can be generally categorised as laboratory serum tests and radiologic tests.
2.3.1. Liver biopsy and histopathological staging system

Liver biopsy remains the reference standard for fibrosis assessment. Percutaneous liver biopsy is the simplest, quickest and the most commonly performed approach in clinical practice. It is an invasive procedure with around 1% risk of serious complication (such as significant bleeding or haemobilia) and less than 0.01% risk of death (60, 61). Moderate pain occurs in 20% to 40% of patients after the procedure (62). For these reasons, liver biopsy is more commonly performed for diagnosis purpose rather than for prognosis or monitoring purpose.

Metavir and Ishak stage are the most widely used histopathological staging systems for chronic hepatitis C. They semi-quantitatively stage liver fibrosis severity according to the fibrosis accumulation and architectural distortion within liver parenchyma. Metavir staging system categorises fibrosis severity into five stages: F0, no fibrosis; F1, portal fibrosis without septa; F2, few septa; F3, numerous septa without cirrhosis; F4, cirrhosis (63). Ishak staging system categorises liver fibrosis severity into seven stages: Stage 0: no fibrosis; Stage 1: fibrous expansion of some portal areas, with or without short fibrous septa; Stage 2: fibrous expansion of most portal areas, with or without short fibrous septa; Stage 3: fibrous expansion of most portal areas with occasional portal to portal bridging; Stage 4: Fibrous expansion of portal areas with marked bridging portal to portal as well as portal to central; Stage 5: Marked bridging with occasional nodules; Stage 6: Probable or definite cirrhosis (64).
Sampling error is the major limitation because that liver biopsy specimen only represents a very small portion (approximately 1/50,000) of the entire liver. A study using liver biopsies with a length ≥15 mm and found that 33% CHC patients had a difference of at least one fibrosis stage between the right and left liver lobes (65). Another study obtained two separate biopsy samples for CHC patients and found 38% patients had a fibrosis stage difference ≥1 and 21% patients had a difference ≥ 2 (66). Moreover, studies showed that smaller biopsy sample size was more likely to cause underestimation of fibrosis severity (67). This suggested that fibrosis accumulation was not entirely evenly distributed through liver parenchyma and a sufficient biopsy sample size was essential to reduce sampling error. It was recommended that a liver biopsies with a length of at least 20 mm and more than 11 complete portal tracts were required to generate reliable results for histopathological staging while a significant proportion of liver biopsies didn’t fulfil this criteria in real clinical practice (67, 68). Observer variability is another concern for histopathological staging system. Several studies found moderate to good agreement for both histopathological scoring system with intra-variability and inter-variability of 0%-4% and 6%-10% respectively (63, 65, 69). The agreement can be further improved by high experience and consensus reading (70).

### 2.3.2. Digital image analysis of liver biopsy

Digital image analysis of liver biopsy was first developed in early 1990’s (71, 72). It is a computer-based procedure including scanning liver biopsies into digital format and measuring the areas of both collagen and remaining liver tissues. The result is presented as the proportion of the area of the biopsy occupied by collagen [collagen proportional area (CPA)] (73). Unlike traditional
histopathological staging systems which mainly depend on the degree of architectural distortion, digital image analysis precisely measures the amount of fibrosis as a continuous variable. Moreover, as a computer-based technology, CPA was more reproducible due to minimised inter and intra-observer variability.

During the past two decades, the potential clinical usage of CPA as an alternative measurement of liver fibrosis is gaining increasingly attention worldwide. Evidence showed that CPA was well correlated with histopathological stages with correlation coefficient of 0.72-0.84 (74-77). The mean CPA value was estimated to be 2%, 3%, 6%, 15% and 25% respectively for each Metavir stage from F0 to F4 (78). A cut point of 6% achieved a sensitivity of 78% and specificity of 80% to detect significant fibrosis (Ishak stage ≥2) and a cut point of 9% achieved a sensitivity of 78% and specificity of 88% to detect cirrhosis (Ishak stage 5 or 6) (76). A wide range of CPA values within each histopathological stage and considerate overlaps within histopathological stages were also observed. This indicated that CPA was a completely different measurement from histopathological staging system and it provided extra important information concerning liver fibrosis severity. As a continuous measurement, CPA had higher sensitivity to identify fibrosis progression or regression compared to that of histopathological staging system (79, 80). CPA also had better correlation with fibrosis serum markers, with hyaluronic acid (HA) and prothrombin time to be the most robust predictors for CPA (75).

Several studies evaluated the ability of CPA to predict outcomes in various clinical settings. CPA was significantly correlated with hepatic venous pressure
gradient (HVPG) with correlation coefficient of 0.61. A cut point of 12.5% had a sensitivity of 78% and a specificity of 86% to predict clinically significant portal hypertension (HVPG ≥ 10 mmHg) (76). Among cirrhotic patients, CPA achieved an Area Under Receiver Operating Characteristic (AUROC) of 0.91 to predict decompensation at the time of liver biopsy and was an independent predictor for future decompensation (81). CPA also had excellent predictive ability for liver decompensation among post-transplantation patients with recurrent HCV infection (82, 83). Moreover, fibrosis progression rate assessed using CPA was a better predictor of clinical outcomes than that using Ishak stage (84).

Although sharing the same principle, various technologies and methods have been used to calculate CPA among studies. Masson’s trichrome is a routine staining method for liver fibrosis and this was used in studies by O’Brien et al (77), Lazzarini et al (85) and Sethasine et al (86). Other studies have used sirius red staining based on its specific binding to collagen (68, 75, 87-89). Different image capture and analysis technologies were also applied. In some studies, biopsy images were captured by a black and white video camera that coupled to a optic microscope and the area of collagen was determined according to grey scare (74, 78, 80). Other studies used colour image for analysis and determined the area of collagen according to hue value and colour saturation (76, 90). A wide range of magnifications of image capturing was found among studies and it varied from 1X by Calvaruso et al (76) to 100X by Bedossa et al (78). Therefore there is a need to standardise each step of CPA measurement and further widen the clinical use of CPA.
2.3.3. Non-invasive assessment of liver fibrosis

Due to the limitation of liver biopsy, non-invasive methods for liver fibrosis assessment have been brought into wide interest. The primary usage of non-invasive tests is to detect the severity of liver fibrosis using histological stage of liver biopsy as a reference standard. Moreover, non-invasive test can be further used to monitor liver fibrosis progression or regression, to assess antiviral treatment efficacy and to predict clinical outcomes.

The ideal non-invasive method should be simple, inexpensive, readily available, reproductive and accurate. The currently developed non-invasive methods can be generally categorised into serum tests and radiologic tests. Serum markers have advantages because of its low cost, wide availability and high reproducibility. However, none of the proposed serum markers is liver specific. Factors other than liver fibrosis severity, such as age, gender and comorbidities have the potential to influence the test results of serum markers and thus reduce their diagnostic accuracy. In contrast, radiologic tests directly evaluate the characters of liver but they are limited by their relatively high cost, expertise requirement and certain technical failure rate.

2.3.4. Serum tests

A number of serum markers that directly involved in the process of fibrogenesis and fibrolysis have been identified. Those serum markers are known as direct serum markers and directly reflect the extracellular matrix (ECM) turn over. Direct serum markers include: glycoproteins, such as HA, laminin, and YKL-40; the collagens fragments, such as procollagen III N-peptide (PIIINP) and type IV
collagen; collagenases and their inhibitors, such as matrix metalloproteases (MMP) and TIMPs. Additionally, a number of serum models were developed in the last 15 years using a selection of serum makers including both direct serum markers and routine serum tests, such as: platelet count, prothrombin time, international normalized ratio (INR), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), bilirubin and albumin.

2.3.4.1. HA

HA is a glycosaminoglycan mainly synthesized by hepatic HSC and degraded by the liver sinusoidal cells. It was suggested that both the increase of HA production and decrease of HA degradation contributed to serum HA increase in patients with liver fibrosis. HA had a mean value of 27-35 ng/ml, ranged from 0 to 63-79 ng/ml among healthy individuals (91, 92). HA showed a close association with liver fibrosis with a correlation coefficient of 0.50-0.57 (93-95). Several studies tested the predictive performance of HA for liver fibrosis and suggested that a HA level less than 60 ng/ml had good accuracy to rule out patients with significant fibrosis or cirrhosis (95-97). Other studies showed that a higher cut point of HA level of 100-237 ng/ml could be used to identify the present of cirrhosis (96, 98). Different assays used to quantify HA in studies may result in slightly different cut point and diagnostic accuracy of HA (99). Moreover, HA level was showed to correlate with the outcome of antiviral treatment for chronic hepatitis C, the development of portal hypertension in patients with chronic liver disease (100) and the development of decompensation in cirrhotic patients (101).
2.3.4.2. PIIINP and type IV collagen

PIIINP is a cleavage product of maturing type III procollagen, produced by activated HSC (102). A large number of early studies observed a significant increase of PIIINP in patients with liver fibrosis. PIIINP correlated with fibrosis stage with correlation coefficient of 0.53-0.62 (94, 103). The AUROC of PIIINP was 0.73 to diagnose cirrhosis and was 0.65-0.69 to diagnose significant fibrosis (104, 105). Type IV collagen is one of the major components of the newly formed basement membranes in fibrotic liver and mainly deposits in perisinusoidal space, which can lead to capillarisation of sinusoid wall. Type IV collagen was correlated with fibrosis stage with correlation coefficient of 0.24-0.45 and had an AUROC of 0.83 to detect advanced liver fibrosis (103, 106).

2.3.4.3. MMPs and TIMPs

MMPs form a family of enzymes that mediates the degradation of extracellular matrix proteins (107, 108). MMPs together with their inhibitors (TIMPs) play an important role in the physiological process of tissue remodelling and wound healing, and in several pathological conditions including liver fibrosis (94). MMP-2 is type IV collagenase and is secreted by activated HSC cells (109). Several studies showed that MMP-2 level positively correlated with liver fibrosis stages with correlation coefficient of 0.26-0.28 (94, 110). TIMP-1 as the specific inhibitor for MMP-2 also increased with liver fibrosis stages (111). TIMP-1 showed a slightly better correlation with liver fibrosis stage than MMP-2 with correlation coefficient of 0.30-0.42 (94, 110). Both MMP-2 and TIMP-1 had a potential to predict the severity of liver fibrosis, but the cut points to diagnosis significant fibrosis or cirrhosis were ill defined due to different assay used
among studies (111). MMP-9 as another type IV collagenase is secreted by activated Kupffer cells (112). Evidence suggested that MMP-9 significantly decreased in patients with liver fibrosis or cirrhosis but its predictive accuracy for liver fibrosis stage was less clear (94, 113). MMP-1 degrades collagen type I, II and III. It showed a significant correlation with inflammation activity but conflict results were found when correlating MMP-1 with fibrosis stages (94, 114).

2.3.4.4. Serum models

The development of serum models that combined a panel of markers substantially improved diagnostic accuracy compared to individual serum markers. A number of serum models have been developed in that last 15 years (table 2.1). Most serum models were able to achieve an AUROC larger than 0.80 to predict significant fibrosis and an AUROC larger than 0.85 to predict cirrhosis (table 2.2). However, considerate overlaps of model values were found between fibrosis stages and these leaded to unacceptable accuracy of serum models to predict individual fibrosis stages. It was commonly observed that the diagnosis accuracy of serum models was satisfying at two extremes of the model values but was poor in the middle. As a result, two cut points were often suggested to detect each fibrosis stage: the lower cut point was used to confidently identify the absence of a certain fibrosis stage and the higher cut point to confidently identify the presence of the same fibrosis stage (table 2.2). In clinical practice, serum models were suggested to be used as screening methods to rule-in or rule-out fibrosis in around 35% patients (115).

Fibrotest and APRI were most extensively validated in external cohorts. A Meta-analysis including 30 studies (3501 CHC patients) found that Fibrotest had an
AUROC of 0.85 to detect significant fibrosis (116). Another Meta-analysis including 40 studies (8739 HCV or HCV/HIV patients) found that APRI achieved an AUROC of 0.77 to detect significant fibrosis and an AUROC of 0.83 to detect cirrhosis (117). Similar AUROC value was found for other commonly used serum models. The AUROC to detect significant fibrosis was 0.75-0.91 for Forns, 0.74-0.85 for FIB-4, 0.74-0.86 for Hepascore, 0.78-0.89 for Fibrometer and 0.77-0.87 for ELF (118). One large study including 1839 patients compared the diagnostic performance of four serum models (Fibrometer, Fibrotest, APRI and Hepascore) and found no significant difference among models to predict significant fibrosis or cirrhosis (119).
Table 2.1. List of serum models

<table>
<thead>
<tr>
<th>Models</th>
<th>components</th>
</tr>
</thead>
<tbody>
<tr>
<td>AST to platelet ratio index (APRI) (120)</td>
<td>AST, and platelet count</td>
</tr>
<tr>
<td>The Forns index(121)</td>
<td>Age, platelet count, GGT, cholesterol levels</td>
</tr>
<tr>
<td>The Hepascore(122)</td>
<td>Age, gender, bilirubin, GGT, HA, α2-macroglobulin</td>
</tr>
<tr>
<td>The FIB-4 score (123)</td>
<td>Age, ALT, AST, platelet count</td>
</tr>
<tr>
<td>The SHASTA index(124)</td>
<td>HA, AST, albumin</td>
</tr>
<tr>
<td>Enhanced Liver Fibrosis score (ELF) (92)</td>
<td>Age, N-terminal propeptide of collagen type III, HA, TIMP-1</td>
</tr>
<tr>
<td>Fibrospect II(125)</td>
<td>HA, TIMP-1, α2-macroglobulin</td>
</tr>
<tr>
<td>Fibrotest (126)</td>
<td>GGT, haptoglobin, bilirubin, apolipoprotein A1, α2-macroglobulin</td>
</tr>
<tr>
<td>Fibroindex(127)</td>
<td>Platelet count, AST, GGT</td>
</tr>
<tr>
<td>Fibrometer(128)</td>
<td>Platelet count, α2-macroglobulin, AST, age, prothrombin index, HA, blood urea nitrogen</td>
</tr>
<tr>
<td>MP3(94)</td>
<td>PIIINP, and MMP-1</td>
</tr>
<tr>
<td>Lok Index (129)</td>
<td>platelet count, AST/ALT ratio, and INR</td>
</tr>
<tr>
<td>Goteborg University Cirrhosis Index (GUCI) (130)</td>
<td>AST, INR, and platelet count</td>
</tr>
<tr>
<td>Fibrosis Probability Index (FPI)(131)</td>
<td>age, past alcohol intake, AST, cholesterol, and HOMA-IR</td>
</tr>
<tr>
<td>Virahep-C (132)</td>
<td>AST, platelet count, alkaline phosphatase, and age</td>
</tr>
<tr>
<td>HALT-C model(133)</td>
<td>hyaluronic acid, TIMP-1, and platelet count</td>
</tr>
</tbody>
</table>

**Note:** HOMA-IR, insulin resistance by the Homeostasis model assessment
Table 2.2. Predictive performance of serum models.

<table>
<thead>
<tr>
<th>Serum models</th>
<th>aetiology</th>
<th>Cut points</th>
<th>interpretation</th>
<th>se</th>
<th>sp</th>
<th>AUROC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APRI (120)</td>
<td>CHC</td>
<td>≤0.5</td>
<td>absence Ishak 3-6</td>
<td>91%</td>
<td>47%</td>
<td>0.80 (0.74-0.87)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;1.5</td>
<td>Present Ishak 3-6</td>
<td>41%</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>≤1.0</td>
<td>absence Ishak 5-6</td>
<td>89%</td>
<td>75%</td>
<td>0.89 (0.84-0.94)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;2.0</td>
<td>Present Ishak 5-6</td>
<td>57%</td>
<td>93%</td>
<td></td>
</tr>
<tr>
<td>Forns index</td>
<td>CHC</td>
<td>&lt;4.2</td>
<td>absence Scheuer 2–4</td>
<td>94%</td>
<td>51%</td>
<td>0.86</td>
</tr>
<tr>
<td>(121)</td>
<td></td>
<td>&gt;6.9</td>
<td>present Scheuer 2–4</td>
<td>30%</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>Hepascore</td>
<td>CHC</td>
<td>0.50</td>
<td>F2-F4</td>
<td>63%</td>
<td>89%</td>
<td>0.85 (0.78–0.93)</td>
</tr>
<tr>
<td>(122)</td>
<td></td>
<td>0.84</td>
<td>F4</td>
<td>71%</td>
<td>89%</td>
<td>0.94 (0.87–1.00)</td>
</tr>
<tr>
<td>FIB-4 score</td>
<td>HCV/HIV</td>
<td>≤0.6</td>
<td>absence Ishak 2-6</td>
<td>92%</td>
<td>23%</td>
<td>0.711</td>
</tr>
<tr>
<td>(123)</td>
<td></td>
<td>&gt;1.0</td>
<td>Present Ishak 2-6</td>
<td>69%</td>
<td>58%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>≤1.45</td>
<td>absence Ishak 4-6</td>
<td>67%</td>
<td>71%</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;3.25</td>
<td>Present Ishak 4-6</td>
<td>23%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>SHASTA index</td>
<td>HCV/HIV</td>
<td>&lt;0.3</td>
<td>absence Ishak 3-6</td>
<td>88%</td>
<td>72%</td>
<td>0.87</td>
</tr>
<tr>
<td>(124)</td>
<td></td>
<td>&gt;0.8</td>
<td>Present Ishak 3-6</td>
<td>15%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>ELF (92)</td>
<td>CLD</td>
<td>&lt;0.10</td>
<td>absence Scheuer 3–4</td>
<td>90%</td>
<td>41%</td>
<td>0.80 (0.76 -0.85)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;0.83</td>
<td>present Scheuer 3–4</td>
<td>19%</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td>Fibrospect II</td>
<td>CHC</td>
<td>0.36</td>
<td>F2-F4</td>
<td>77%</td>
<td>73%</td>
<td>0.83 (0.79–0.88)</td>
</tr>
<tr>
<td>(125)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrotest</td>
<td>CHC</td>
<td>0.48</td>
<td>F2-F4</td>
<td>75%</td>
<td>85%</td>
<td>0.87</td>
</tr>
<tr>
<td>(126)</td>
<td></td>
<td>0.74</td>
<td>F4</td>
<td>63%</td>
<td>84%</td>
<td>0.87</td>
</tr>
<tr>
<td>Fibroindex</td>
<td>CHC</td>
<td>&lt;1.25</td>
<td>absence F2-F3</td>
<td>40%</td>
<td>94%</td>
<td>0.86 (0.81 -0.92)</td>
</tr>
<tr>
<td>(127)</td>
<td></td>
<td>&gt;2.55</td>
<td>Present F2-F3</td>
<td>36%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>Fibrometer</td>
<td>VH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.88 (0.85- 0.92)</td>
</tr>
<tr>
<td>(128)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Continued as Table 2.2

<table>
<thead>
<tr>
<th>Serum models</th>
<th>aetiology</th>
<th>Cut points</th>
<th>interpretation</th>
<th>se</th>
<th>sp</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibroindex (127)</td>
<td>CHC</td>
<td>&lt;1.25</td>
<td>absence F2-F3</td>
<td>40%</td>
<td>94%</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;2.55</td>
<td>Present F2-F3</td>
<td>36%</td>
<td>97%</td>
<td>(0.81-0.92)</td>
</tr>
<tr>
<td>Fibrometer (128)</td>
<td>VH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.85-0.92)</td>
</tr>
<tr>
<td>MP3 (94)</td>
<td>CHC</td>
<td>&lt;0.3</td>
<td>absence F2-F4</td>
<td>65%</td>
<td>85%</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;0.5</td>
<td>Present F2-F4</td>
<td>26%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>Lok Index (129)</td>
<td>CHC</td>
<td>&lt;0.2</td>
<td>absence Ishak 5-6</td>
<td>48%</td>
<td>92%</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;0.5</td>
<td>Present Ishak 5-6</td>
<td>12%</td>
<td>54%</td>
<td>(0.74–0.81)</td>
</tr>
<tr>
<td>GUCI (130)</td>
<td>CHC</td>
<td>1.0</td>
<td>Ishak 5-6</td>
<td>80%</td>
<td>78%</td>
<td>0.85</td>
</tr>
<tr>
<td>Fibrosis Probability</td>
<td>CHC</td>
<td>&lt;0.2</td>
<td>absence F2-F4</td>
<td>96%</td>
<td>44%</td>
<td>0.84</td>
</tr>
<tr>
<td>Index (FPI) (131)</td>
<td></td>
<td>≥0.8</td>
<td>Present F2-F4</td>
<td>43%</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>Virahep-C model (132)</td>
<td>CHC</td>
<td>≤0.22</td>
<td>absence Ishak 3-6</td>
<td>90%</td>
<td>54%</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥0.55</td>
<td>Present Ishak 3-6</td>
<td>51%</td>
<td>90%</td>
<td>(0.80-0.88)</td>
</tr>
<tr>
<td>HALT-C model (133)</td>
<td>CHC</td>
<td>&lt;0.2</td>
<td>absence Ishak 5-6</td>
<td>88%</td>
<td>45%</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;0.5</td>
<td>Present Ishak 5-6</td>
<td>47%</td>
<td>92%</td>
<td>(0.77-0.85)</td>
</tr>
</tbody>
</table>

**Note:** CLD: chronic liver disease; VH: viral hepatitis; ALD: alcoholic liver disease; SF: significant fibrosis.
2.3.5. Radiologic tests

2.3.5.1. Ultrasound-based transient elastography

Ultrasound-based transient elastography is an ultrasound-based technology that measures liver stiffness (134). By using an ultrasonic transducer probe mounted on the axis of a vibrator, a vibration of a low frequency (50 Hz) and amplitude is transmitted into liver and induces an elastic shear wave that propagates through the organ. The velocity of the wave propagation is then measured by the probe. The velocity, as expressed in kilopascals (kPa), is proportionally correlated with liver stiffness: the harder the tissue, the faster the propagation of elastic waves (135). Transient elastography is performed in the right lobe of the liver, through the intercostal space. This technic allows the measurement of liver stiffness using a liver section with approximately 1cm wide and 5 cm long. It is 100 times greater in size than a standard liver biopsy and thus potentially reduces the variability caused by sampling error.

The normal value of transient elastography is not entirely clear due to the lack of a large control group. A study of 429 healthy individuals found the normal value was estimated at 3.3–7.8 kPa in women and 3.8–8.0 kPa in men (136). The diagnosis accuracy of transient elastography has been extensively studied. Transient elastography showed high diagnostic accuracy for cirrhosis with an AUROC larger than 0.90 in most studies. A cut point of 12.5 kPa or more achieved a pooled sensitivity of 87% and pooled specificity of 91% (137). The diagnostic ability was relatively poor for significant fibrosis with an AUROC varied from 0.74 to 0.94 among studies. A cut point of 4 kPa had a sensitivity of 94% and a specificity of 33% to detect significant fibrosis and a cut point of 8.8
kPa achieved a sensitivity of 0.56% and a specificity of 0.91 to detect cirrhosis (138-140). A Meta-analysis including 50 studies reported that the mean AUROC for the diagnosis of significant fibrosis (Metavir F2-F4), severe fibrosis (F3-F4) and cirrhosis (F4) was 0.84, 0.89 and 0.94 respectively (141). The optimal cut point was 7.65 kPa for significant fibrosis and 13.01 kPa for cirrhosis (141). Additionally, the performance of transient elastography was tested for detection of liver fibrosis in patients with recurrent HCV after liver transplantation. It has excellent diagnostic accuracy to identify cirrhosis and better ability to detect significant fibrosis for this cohort compared to pre-transplant patients with chronic hepatitis C (142).

One important limitation of transient elastography is that around 5% of technical failure rate and 15% chance of unreliable results were observed during clinical practice and these were associated with high body mass index, high waist circumference, low operator experience, older age, female sex, hypertension, and type 2 diabetes (143, 144). Moreover, evidence showed that the reproducibility of transient elastography was significantly reduced in patients with steatosis, increased body mass index and lower degrees of hepatic fibrosis (145). Factors other than liver fibrosis such as acute hepatitis, metabolic syndrome also cause transient elastography value to increase (136, 146).

2.3.5.2. Acoustic radiation force impulse imaging

Acoustic radiation force impulse imaging permits evaluation of the elastic properties of a region of interest while performing a real-time B-mode ultrasonography. This technique uses short-duration, high intensity acoustic pulses to produce mechanical excitation and generate localized tissue
displacements in liver. The displacements result in shear wave propagation away from the excitation region and this is tracked by ultrasound correlation-based methods. By measuring the time to peak displacement at each lateral location, the velocity of shear wave propagation can be reproduced (147). Results are expressed in meters per second. The region of interest is chosen under ultrasound guidance thus large blood vessels and liver capsules can be avoided. Moreover, the inclusion in a conventional ultrasound machine facilitates the application in clinical practice.

Several studies validated the diagnostic performance of acoustic radiation force impulse imaging but most of them were limited by small number of patients. One of the largest studies included 321 patient with chronic liver disease found the AUROC of acoustic radiation force impulse imaging was 0.81 for significant fibrosis and 0.88 for cirrhosis (148). Compared with transient elastography, acoustic radiation force impulse imaging had similar diagnostic accuracy for advance fibrosis and cirrhosis but less accuracy for significant fibrosis (147-151). Conflicting results were found concerning the effect of obesity on the diagnostic performance of acoustic radiation force impulse imaging (148, 152). Additionally, no quality criteria for acoustic radiation force impulse imaging measurement are introduced by studies.

2.3.5.3. Real-time shear wave elastography

Real-time shear wave elastography is based on the same principles as ultrasound-based transient elastography and measures the velocity of shear wave propagation in liver tissue (kPa). In contrary to ultrasound-based transient elastography, in real-time shear wave elastography, the shear waves are
generated by focussed ultrasound beams and the speed of shear wave is estimated over a region with the guidance of a B-mode image (153). The result is displayed as the mean and standard deviation of elasticity for a chosen region of interest. As a result, real-time shear wave elastography allows real time measurement of liver elasticity and avoidance of big blood vessel and liver capsule.

The reproducibility of real-time shear wave elastography was tested in one study using 42 healthy individuals. The intra and inter-observer agreement was 0.93-0.95 and 0.88 respectively (153). Another study including 121 chronic hepatitis C patients found that real-time shear wave elastography achieved an AUROC of 0.92 for significant fibrosis (Metavir F2-F4), 0.98 for severe fibrosis (Metavir F3-F4) and 0.98 for cirrhosis (Metavir F4). It has significant better diagnostic accuracy for significant fibrosis than ultrasound-based transient elastography and similar accuracy for advanced fibrosis and cirrhosis (154). No quality criteria for shear wave elastography measurement are used by studies.

### 2.3.5.4. Magnetic resonance elastography

Magnetic resonance elastography is a technique for quantitatively assessment of the mechanical properties of soft tissues and has the potential to noninvasively detect liver fibrosis. Similar to ultrasound-based transient elastography, a probe or electromechanical driver is used to generate a low frequency vibration that pass through the liver. A specialized phase-contrast magnetic resonance imaging sequence is used to image the propagation characteristics of elastic shear waves in the liver. The wave images are then processed and the mean elasticity values measured in the liver are obtained.
and expressed as kPa (155). Compared to ultrasound-based transient elastography, magnetic resonance elastography measures the mean stiffness of the entire liver and thus generates more reliable results. A Meta-analysis including four studies found that the mean AUROC, sensitivity and specificity for magnetic resonance elastography to detect significant fibrosis (F2-F4) was 0.98, 94% and 95% respectively (156). The mean AUROC, sensitivity and specificity to detect severe fibrosis (F3-F4) was 0.98, 92% and 96% respectively (156). A few studies compared the diagnostic performance of magnetic resonance elastography with ultrasound-based transient elastography and found better accuracy and higher technical success rate of magnetic resonance elastography (157).

### 2.3.5.5. Diffusion-weighted magnetic resonance imaging

Diffusion-weighted magnetic resonance imaging measures the apparent diffusion coefficient of water, a parameter that is dependent on the tissue structure (158). Studies showed that the apparent diffusion coefficient value significantly decreased with fibrosis stages due to the restriction of water in fibrotic liver tissues (158, 159). A Meta-analysis found that diffusion-weighted magnetic resonance imaging achieved less diagnostic accuracy compared to magnetic resonance elastography (156). It had an AUROC, sensitivity and specificity of 0.83, 77% and 78% respectively to detect significant fibrosis and 0.86, 72% and 84% respectively to detect advanced fibrosis (156). Other study compared diffusion-weighted magnetic resonance with ultrasound-based transient elastography and serum markers (FibroTest, APRI, Forns index, and hyaluronate) and suggested that it has the best diagnostic performance compared to other non-invasive tests to determine advanced fibrosis (158).
2.3.6. Combining tests

The combination of several non-invasive tests may improve the diagnostic accuracy for liver fibrosis and reduce the number of biopsies needed. One study evaluated the performance of transient elastography (FibroScan) and Fibrotest and found that FibroScan and Fibrotest agreed in 70%-80% of patients. The agreed results in consist with liver biopsy examination for significant fibrosis (Metavir F2-4) in 84% of cases, severe fibrosis (Metavir F3-4) in 95%, and cirrhosis (Metavir F4) in 94%. This suggested that by combining FibroScan and Fibrotest, liver biopsy can be avoided in 77% of the patients (160). Another group proposed two algorithms combining Fibrotest and APRI (SAFE biopsy) to detect significant fibrosis and cirrhosis respectively (161). SAFE biopsy was validated in 2035 patients with chronic hepatitis C and achieved an accuracy of 90.1% to identify significant fibrosis and 92.5% to identify cirrhosis. It obviated 46.5% and 81.5% of liver biopsy respectively for each purpose (162). The combination of FibroScan and Fibrotest was then compared with SAFE biopsy in 314 CHC patients. The two combinations achieved similar ability to diagnose significant fibrosis and cirrhosis with an accuracy of 88% to 97%. However, the number of liver biopsies avoided was significantly higher in FibroScan and Fibrotest combination than SAFE biopsy for significant fibrosis (163).
2.4. OUTCOME PREDICTION FOR CHRONIC HEPATITIS C

The major clinical outcomes of chronic hepatitis C are liver decompensation, HCC development and death. As chronic hepatitis C is a slowly progressive disease with a long term asymptomatic phase, the potential risk to develop adverse clinical outcomes for CHC patients is usually uncertain. The huge variability of disease progression rate among individuals further adds the difficulty to accurately predict outcome for each patient. Thus methods that can accurately predict clinical outcomes and identify patients with higher risk are of great clinical significance. Cirrhosis is generally recognised as the most advanced stage of chronic liver disease where liver related morbidity and mortality start to occur. Numerous efforts have been made to illustrate the natural history of cirrhosis and to identify predictors for clinical outcomes among cirrhotic patients. Meanwhile, the outcome prediction for non-cirrhotic patients is gaining increasingly interest and this will help clinicians to identify patients with higher risk at earlier phase of the disease.

2.4.1. Outcome predictors for cirrhotic patients

The diagnosis of cirrhosis is primarily relies on abnormal nodular formation within the liver. It is widely accepted that cirrhosis is the most advanced stage of chronic liver disease with dramatically increased risk of the development of liver decompensation, HCC and liver related death. Increasing evidence suggested that cirrhosis was not only a static histopathological diagnosis but a dynamic stage where further progressions in anatomical, hemodynamic and biochemical status took place and eventually leaded to death (164). Further risk stratification
for cirrhotic patients is essential for individualised management including antiviral treatment, liver complication surveillance and liver transplantation.

### 2.4.1.1. Clinical manifestations

Classification of cirrhosis into a compensated and a decompensated stage is the simplest and most reproducible way to stratify risk in clinical practice. It is well documented that patients with decompensated cirrhosis had a markedly shorter life expectancy than those with compensated cirrhosis (165, 166). The median survival time was longer than 12 years for compensated cirrhotic patients and was about two years for those with decompensated cirrhosis (167). Additionally, a systemic review, proposed a more specific staging system that categorised cirrhosis into four stages based on the occurrence of esophageal varices and variceal bleeding (167). Stage 1 was characterised by compensated cirrhosis with no varices; stage 2 was characterized by compensated cirrhosis with esophageal varices; stage 3 was characterized by ascites without esophageal bleeding and stage 4 was characterised by variceal bleeding with or without ascites. The annual mortality rate for each stage was estimated to be 1%, 3.4%, 20% and 54% respectively. Following studies found that appearance of varices (stage 2) was significantly associated with higher risk of decompensation and death among compensated cirrhotic patients (165, 168, 169). Conflicting results were found concerning the risk stratification ability of the occurrence of variceal bleeding among decompensated patients (165). Additionally, infection and renal failure were associated with four fold and seven fold increase of mortality rate in cirrhotic patients (170, 171).
2.4.1.2. Extra histopathological features and image analysis

Additional histological features that are not incorporated into the classical fibrosis staging systems may also have important prognostic implications and thus have the ability to sub-stage cirrhosis. One study showed that nodule size and fibrous septa thickness were independent predictors for the presence of clinically significant portal hypertension among cirrhotic patients (172). Kumar et al proposed a sub-classification score according to nodule size (large=1, medium=2, small=3) and septal thickness (thick=3, medium=2, thin=1). Two subcategories were devised based on the total score: category A: score 1-3 and category B: score 4-6 and this was significantly associated with the presence of clinically significant portal hypertension (173). Laennec staging system sub-categorised cirrhosis into three stages: 4A: visible nodules, thin septa and up to one broad septum present; 4B: at least two broad septa, but no very broad septa and less than half of biopsy length composed of small nodules; 4C: At least one very broad septum or more than half of biopsy length composed of small nodules (174). Laennec staging system was significantly associated with the presence of clinically significant portal hypertension and the risk of adverse clinical outcomes in cirrhotic patients (174-176).

Image analysis quantitatively measures the amount of collagen accumulation in liver biopsy. A wide range of CPA value was observed among cirrhotic patients, ranging from 1.4% to 65% (77). This suggested that CPA had the potential to sub-stratify risk of adverse clinical outcomes for cirrhotic patients. Evidence showed that CPA was closely correlated with HVPG and was superior to all the other cirrhosis sub-classification systems to predict clinical decompensation among cirrhotic patients (76, 81).
2.4.1.3. Hepatic venous pressure gradients

As portal hypertension is the underlying reason for most complications of cirrhosis, HVPG, as an estimation of portal hypertension, became the gold standard for risk prediction among cirrhotic patients. HVPG is measured as the difference in pressure between the portal vein and inferior vena cava and is normally between 1 to 5 mmHg. HVPG value between 5 to 9 mmHg is recognised as sub-clinical portal hypertension. Clinical significant portal hypertension is defined as an HVPG ≥ 10 mmHg, at which point esophageal varices may develop (177). Among compensated cirrhotic patients, those patients with an HVPG ≥ 10 mmHg had a significantly higher risk of the development of HCC, decompensation and death (165, 178-180). Those patients with an HVPG ≥ 12 mmHg were at risk of variceal bleeding (181, 182). Additionally, for cirrhotic patients under pharmacologic treatment, a decrease in HVPG ≥ 20% or to 12 mm Hg or less was associated with a marked reduction in the long-term risk of developing portal hypertension associated complications and improved survival (183). An HVPG ≥ 20 mmHg during acute variceal bleeding was associated with higher risk of five day treatment failure and lower one year survival (184). Among patient with ascites, those with a HVPG ≥ 30 mmHg had significantly higher risk of develop spontaneous bacterial peritonitis (185). Used as a continuous variable, HVPG also showed independent predictive value on survival in addition to the MELD score among cirrhotic patients (186).
2.4.1.4. Serum markers and models

A large number of serum markers were significantly associated with death among cirrhotic patients and those included: albumin, bilirubin, prothrombin time, GGT, creatinine, ALP, sodium, AST, HA, blood urea nitrogen, and tumor necrosis factors. Among them, albumin, bilirubin, prothrombin time were the most commonly found predictors for death (167). Child-Pugh score and Model for End-stage Liver Disease (MELD) score were the most widely used models to predict death for cirrhotic patients.

The Child-Pugh score was designed to stratify the risk of portacaval shunt surgery in patients with cirrhosis (187). Variables included in Child-Pugh score are ascites, bilirubin, albumin, prothrombin time and encephalopathy (table 2.3). The score ranges from 5 to 15. Patients with a score of 5 or 6 have Child-Pugh class A cirrhosis (well-compensated cirrhosis), those with a score of 7 to 9 have Child-Pugh class B cirrhosis (significant functional compromise), and those with a score of 10 to 15 have Child-Pugh class C cirrhosis (decompensated cirrhosis). A systemic review of 118 studies found that Child-Pugh score was the most robust predictor of death for cirrhotic patients (167). The one year survival probability for patients with Child-Pugh class A, B and C was approximately 95%, 80% and 45% respectively (167). Child-Pugh score significantly associated with liver related death and HCC development among compensated patients (168). Among patients with ascites, Child-Pugh score was an independent predictor for death (188, 189). One limitation of Child-Pugh score was that it included subjective variables such as ascites and encephalopathy and this could potentially lead to higher inter and intra-variability.
Table 2.3. Child-Turcotte-Pugh score.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Point assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Ascites</td>
<td>Absent</td>
</tr>
<tr>
<td>Bilirubin (mmol/L)</td>
<td>&lt;34.2</td>
</tr>
<tr>
<td>Albumin (g/L)</td>
<td>&gt;35</td>
</tr>
<tr>
<td>Prothrombin time</td>
<td></td>
</tr>
<tr>
<td>Seconds over control</td>
<td>&lt;4</td>
</tr>
<tr>
<td>INR</td>
<td>&lt;1.7</td>
</tr>
<tr>
<td>Encephalopathy</td>
<td>None</td>
</tr>
</tbody>
</table>
MELD score was first developed to predict short term clinical outcomes following the transjugular intrahepatic portosystemic shunt procedure (190). It is calculated as: $10 \times ((0.957 \times \ln(\text{Creatinine (mmol/L)/88.4})) + (0.378 \times \ln(\text{Bilirubin (mmol/L)/17.1})) + (1.12 \times \ln(\text{INR}))) + 6.43$. As MELD score does not include subjective variables, it is considered more reproducible than Child-Pugh score and is now widely used for prioritisation of organ allocation in patients on the liver transplantation waitlist. Evidence showed that MELD score had excellent accuracy in predicting three month mortality for patients with end stage liver disease with an AUROC of 0.78-0.87 (191, 192). The application of MELD score in transplantation settings achieved significantly reduced waiting list registration rate and reduced waiting list mortality rate compared to that in pre-MELD era (192-194). In addition to organ allocation, the clinical applications of MELD score include prognosis for acute variceal bleeding, infection, alcoholic hepatitis, surgical resection of hepatocellular carcinoma, placement of transjugular intrahepatic portosystemic shunt, fulminant hepatic failure and renal failure (195). A number of studies tested the predictive ability of MELD in patients with different disease severity. MELD score showed satisfactory accuracy to predict short or mediate term survival for cirrhotic patients (196) and for only decompensated patients (165). However, its predictive accuracy decreased significantly among compensated cirrhotic patients, or to predict long term outcomes (167, 197).

Several attempts were made to develop serum models to predict HCC development among cirrhotic patients (198-200). One study developed a model that categorised patients into 16 risk groups of HCC development according to alpha-fetoprotein, age, gender and platelet count. The five and ten year HCC
risk varied from 9% to 64% and from 21% to 93% among the groups (199). Another study developed a model including prothrombin activity, age, platelet count and presence of anti-HCV antibodies. The model value varied from 0 to 4.72. Using a cut point of 2.33, patients were categorised into low and high risk group with four year HCC risk of 2.3% and 30.1% respectively (198). The third model included both serum markers and clinical manifestations (age, gender, esophageal varices, prothrombin time, alpha-fetoprotein and anti-HCV antibody) and stratified cirrhotic patients into two risk groups with three year HCC free survival of 100% and 76% respectively (200).

2.4.1.5. Ultrasound-based transient elastography

Transient elastography has the potential to further stratify cirrhotic patients into different risk levels. A wide range of transient elastography value among cirrhotic patients was observed, ranging from 5 kPa to 80 kPa (201). A close correlation between transient elastography and HVPG was found in a number of studies with a correlation coefficient ranging from 0.55 to 0.86 (202-206). However, the correlation was less strong for those patients with an HVPG ≥ 10 mmHg and those with an HVPG ≥ 12 mmHg, with a correlation coefficient of 0.59 and 0.37 respectively (205). The AUROC of transient elastography to detect clinical significant portal hypertension (HVPG ≥ 10 mmHg) was 0.76-0.99 and the suggested cut point varied from 13.6 kPa to 34.9 kPa (202-206). The big variability of the suggested cut point was partly due to the different aetiologies of chronic liver disease included among studies. One study found that the cut point of transient elastography to detect clinical significant portal hypertension was higher in patients with alcoholic liver disease than that in those with chronic hepatitis C (203). The close correlation between transient
elastography and HVPG suggested that transient elastography was a promising tool to detect the presence of esophageal varices. Studies found that the AUROC of transient elastography to detect esophageal varices was 0.74-0.85 (204-209). Using a cut point of 13.9 kPa, transient elastography achieved a sensitivity of 95%-96% and a specificity of 39%-43% (206, 207). The higher sensitivity of transient elastography allowed a satisfying accuracy to rule out those patients without esophageal varices. However, as several studies suggested, it could not replace endoscopy for esophageal varices screening. In addition to predicting esophageal varices, a transient elastography value ≥ 40 kPa was significantly associated with higher risk of liver decompensation for cirrhotic patients with HIV/HCV co-infection (210).

2.4.2. Outcome prediction beyond cirrhotic patients

Outcome prediction for non-cirrhotic patients is more difficult due to the long asymptomatic phase of disease before the development of cirrhosis. Histopathological staging system of liver fibrosis remains the gold standard for patient prognosis as liver fibrosis accumulation is the major factor that drives the disease progression. However, due to the invasive manner of liver biopsy, frequent or multiple applications of liver biopsy is poorly tolerated. Non-invasive tests for fibrosis assessment such as serum models and transient elastography have also been validated to predict clinical outcomes for CHC patients. Moreover, a few studies aimed to develop a serum model to directly predict outcomes for patients with advance fibrosis.
2.4.2.1. Histopathological fibrosis stage

It is a well-known fact that fibrosis accumulates within liver tissue along with the progression of chronic hepatitis C and the majority of liver related morbidity and mortality occurs after the development of cirrhosis. For these reasons, histopathological fibrosis stage remains the gold standard to predict clinical outcomes. A number of studies have demonstrated that the risk of adverse clinical outcomes increased with fibrosis severity (211-213). Other studies aimed to evaluate the rate of adverse clinical outcomes for each histopathological fibrosis stage. One study including 1050 CHC patients with significant liver fibrosis found the six year cumulative incidence of first liver related outcome was 5.6% for Ishak stage 2, 16.1% for stage 3, 19.3% for stage 4, 37.8% for stage 5, and 49.3% for stage 6 (214). Another study included 131 patients with Ishak stage 4 to 6 and found the five year risk of liver related morbidity and mortality was 33% and 22% respectively (215). While the rate of adverse clinical outcomes is well described for CHC patients with severe fibrosis, further evidence is needed to illustrate the long terms risk for those patients with less severe liver fibrosis.

2.4.2.2. Serum markers

A few serum fibrosis models showed the correlation with clinical outcomes. A recent Meta-analysis reported that Fibrotest, APRI and FIB-4 had more than one prognostic validations among patients with mixed aetiologies of chronic liver disease: Fibrotest (4 studies; 2,396 patients), APRI (5 studies; 2,422 patients), and FIB-4 (3 studies; 1,184 patients) (216). The AUROC was 0.88 for Fibrotest, 0.73 for FIB-4, and 0.66 for APRI to predict liver related survival (216).
Among patients with chronic hepatitis C, Fibrotest achieved an AUROC of 0.96 to predict liver decompensation and an AUROC of 0.96 to predict liver related death and this was significantly higher than that of APRI and Forn’s index (212). Hepascore was validated in a study with 406 CHC patients and achieved an AUROC of 0.86 to predict liver related death (217). Another study followed 457 patients with mixed aetiology of chronic liver disease and found ELF achieved an AUROC of 0.86 to predict liver related incidence (liver related death or decompensation) (211).

Three studies followed patients with advanced liver fibrosis (Ishak 3-6) for up to five years and developed serum models to directly predict clinical outcomes (218-220). Bilirubin, albumin, INR and YKL-40 levels were included in one model and successfully divided patients into low, median and high risk groups with the four year risk of adverse outcome of 8%, 30% and 65% respectively (218). Another model included only routine serum tests (AST/ALT, albumin, platelet count and bilirubin) and divided patients into four risk groups (219). The third model including both serum tests and clinical features (age, race, platelet count, alkaline phosphatase, esophageal varices and smoking) was developed to specifically predict HCC development (220). These models showed the ability to categorise patients into significantly different risk levels of adverse clinical outcomes. However, accuracy assessment including AUROC, sensitivity and specificity of these models were not reported. Additionally, one study followed patients with all spectrum of liver fibrosis severity for a mean of seven years and identified four serum markers that independently correlated with HCC development: age, alpha-fetoprotein level, gamma globulin and platelet count.
The model (HCC-4) was developed accordingly and achieved an AUROC of 0.802 for HCC prediction (221).

This suggested a promising role of serum markers for disease prognosis in chronic hepatitis C. Further prospective studies are of great interest to validate and compare the prognostic performance of the proposed serum models and thus the most appropriate one can be identified and used as a standard laboratory test in clinical practice in the future (222).

### 2.4.2.3. Transient elastography

The predictive ability for transient elastography has been validated in many studies. A recent Meta-analysis including 17 studies (7058 patients) found that transient elastography was significantly correlated with the development of liver decompensation, HCC, death and composite of these outcomes for patients with chronic liver disease with a risk ratio of 1.07, 1.11, 1.22 and 1.32 respectively (223). However, the corresponding AUROC value and cut point to predict each outcome was not presented in this study. Two studies evaluated the prognostic accuracy of transient elastography for chronic hepatitis C patients. One study followed 1457 patients and found that transient elastography achieved an AUROC of 0.82 to predict five year overall survival (213). The combination of transient elastography and Fibrotest achieved a significantly higher AUROC value of 0.90 compared to that of transient elastography alone (213). Another study categorised patients into five risk groups according to transient elastography (<10 kPa, 10-15 kPa, 15-20 kPa, 20-25 kPa and >25 kPa) (224). The hazard ratio for HCC development was 16.7, 20.9, 25.6 and 45.5 respectively for the last four groups compared to the
first group (224). These suggested that transient elastography was significantly correlated with mortality, decompensation and HCC development. However, its prognostic accuracy and the optimum cut point to identify patients with high risk is yet to be further evaluated.
2.5. SUMMARY STATEMENT

Chronic hepatitis C is a slowly progressive disease with a long term asymptomatic phase in its early stage. Liver fibrosis accumulates along with the progression of chronic hepatitis C infection and eventually leads to cirrhosis. The consequences of chronic hepatitis C include liver decompensation, HCC and liver related death and they almost exclusively occur after the development of cirrhosis. The determination of fibrosis severity is of great importance for patient prognosis and management.

Histopathological stage system of liver biopsy is the gold standard for fibrosis assessment. Metavir and Ishak staging system are most widely used. They semi-quantitatively categorised fibrosis severity into five or seven stages according to fibrosis accumulation and architecture distortion within the liver. However, liver biopsy has well documented limitations such as sampling error, high cost and risk of serious complications and death (60-62, 66). Furthermore, liver biopsy is inconvenient and not widely accessible to either patients or physicians. Therefore, simple, accurate and validated alternatives to liver biopsy are currently required.

Non-invasive methods for liver fibrosis assessment were widely developed in the last two decades. The non-invasive methods can be generally categorized in to serum tests and radiologic tests. Serum models have advantages because of its low cost, wide availability and high reproducibility. However, the reference standard used to create serum models (histopathological stage system) has limitations, such as inter and intra-observer variability and a limited number of categories used to describe the severity of liver fibrosis. Evidence showed that
other histopathological features such as nodular size and septa thickness can provide extra information for fibrosis severity assessment and patient prognosis (172). These limitations may potentially reduce the accuracy of developed serum models to detect fibrosis severity.

Image analysis of liver biopsy is a newly developed method to assess the severity of liver fibrosis. Image analysis is based on quantitatively measurement of the area of collagen as well as the area of remained liver tissues. Result is expressed as proportion of the area of the biopsy occupied by collagen (CPA). CPA has the advantage in its quantitative manner of fibrosis measurement and its high reproducibility. Compared to histopathological stage, CPA has higher sensitivity to detect fibrosis progression and regression and has better correlation with serum markers (75, 79, 80). CPA also has the potential to sub-stage cirrhosis and independently predict outcomes for cirrhotic patients (76, 81). However, existing studies only evaluated the prognosis ability of CPA for cirrhotic patients or those patients after liver transplantation. The predictive ability of CPA for patients with the whole spectrum of fibrosis severity is unclear.

The natural history and prognosis factors for cirrhotic patients have been extensively studied. However, the outcome prediction for CHC patients with less severe fibrosis is still a challenge in clinical management. As the risk for adverse clinical outcomes closely associated with liver fibrosis severity, histopathological stage remains the gold standard for patient prognosis. Non-invasive methods such as serum models for fibrosis assessment and transient elastography have also been validated to predict clinical outcomes. Few efforts have been made to developed serum models to directly predict clinical
outcomes and most of them only targeted for those patients with advanced fibrosis or cirrhosis.

The primary goal of this project was to develop and validate new methods to detect liver fibrosis severity. To fulfil this goal, the optimum methods for CPA measurement was determined and validated. The predictive ability of CPA was evaluated and compared with Metavir stage. A serum model was developed to predict fibrosis severity using CPA as the reference standard. The second goal of this project was to develop serum models to directly predict clinical outcomes for CHC patients.
2.6. AIMS OF THE THESIS

Aim 1: To evaluate long term outcomes for CHC patients stratified by all Metavir fibrosis stages (chapter 3).

Aim 2: To compare CPA obtained using different staining methods, magnifications and biopsy sizes and to determine the optimum method of CPA measurement (chapter 4)

Aim 3: To compare the ability of the CPA with Metavir stage to predict long term liver related morbidity and mortality for CHC patients (chapter 5).

Aim 4: To test the correlation of serum markers and CPA and to build serum model to predict CPA values (chapter 6).

Aim 5: To develop serum models that directly predict the risk of liver related death, HCC and decompensation for CHC patients (chapter 7).
CHAPTER 3

CLINICAL OUTCOMES OF CHRONIC HEPATITIS C PATIENTS RELATED TO BASELINE LIVER FIBROSIS STAGE – A HOSPITAL BASED LINKAGE STUDY

3.1. ABSTRACT

3.1.1. Background and aims

Rates of long term clinical outcomes of chronic hepatitis C in patients with none, mild or severe liver fibrosis are required to determine benefits of anti-viral therapies. This study evaluated long term outcomes for chronic hepatitis C stratified by all Metavir fibrosis stages.

3.1.2. Methods

Clinical outcomes were determined using population based data-linkage methodology for 880 hepatitis C patients who had a liver biopsy performed from 1992 to 2012.

3.1.3. Results

During 9386 person-years of follow-up, 28 patients developed hepatocellular carcinoma, 58 developed liver decompensation and 122 died or underwent liver transplantation. There was no significant difference in liver related death for those with F0-F2 with an 18 year survival probability > 94%. Hazard ratio of
liver related death for F3 compared to F0-F2 was 4.24 (P=0.003), with no significant difference in the first 13 year follow up. The 15 year decompensation free survival for F0, F1 and F2 was 100%, 96% and 94% respectively and for hepatocellular carcinoma free survival was 100%, 99% and 98%. Hazard ratio of liver complication (hepatocellular carcinoma or decompensation) free survival for F3 compared to F0-F2 was 3.22 (P=0.001), with no significant difference during the first seven year follow up. F4 had significantly higher risk of liver related death, decompensation and hepatocellular carcinoma than F3 (p<0.001).

3.1.4. Conclusions

Chronic hepatitis C patients with F2 or less had few liver complications after 15 years. For F3 patients the significant increase in liver related death occurred after 13 years and for liver complications after seven years.
3.2. INTRODUCTION

Hepatitis C virus (HCV) infection affects about 180 million people worldwide and predisposes these individuals to complications of cirrhosis and death (225). Patients with chronic hepatitis C (CHC) after five to seven years of follow up had a three times higher risk of overall death and a 17 times higher risk of liver related death than the general population (226, 227).

Descriptions of the natural history of HCV infection vary greatly and this in part is due to constraints of study design as a consequence of the uncertain time of acquisition of infection and the prolonged course of disease. Nevertheless, a number of well documented host-related factors including age, gender, race and alcohol consumption influence the progression of CHC and add to the variability of individual disease progression rates (23, 26, 228). Viral factors are mostly associated with treatment outcomes rather than progression of CHC (23). The risk of liver related morbidity and mortality was minimal in the first two decades after acquisition of HCV infection but divergent results were found beyond this time. After 25 years follow up the rate of hepatocellular carcinoma (HCC) development and liver related death in young German women infected with HCV following anti-D globulin use was 1.5% and 0.5% respectively (12). Another study after 31 years follow up reported worse outcomes with HCC development in 4.7% and liver related death in 5.3% (20).

In clinical practice the vast majority of CHC patients present after an unknown time of infection so that the severity of liver disease may potentially vary from none to severe. Therefore liver biopsy and more recently non-invasive markers of liver fibrosis have been used to determine the severity of liver fibrosis at a
given time point and allow prediction of the risk of liver related adverse outcomes and determine the timing of treatment. The rate of adverse clinical outcomes is well described for those HCV patients with severe fibrosis but data is lacking for those with less severe liver fibrosis.

The aim of this study was to determine the long term clinical outcome of CHC patients with all stages of liver fibrosis in real world using the linked data from the Western Australia Department of Health. The Western Australia Data Linkage Unit is a validated population-based data linkage system that links multiple health related datasets including cancer registrations, inpatient hospital morbidity and mortality records dating back to 1982, 1970 and 1969 respectively (229). The Hospital Morbidity Data System has 100% coverage of data for all hospitals admissions in Western Australia and has been widely used in cohort and population based studies (230-232). This study will provide important new information that will better guide individual evaluation and management of CHC patients.
3.3. METHODS

3.3.1. Patients

1033 CHC patients who had a liver biopsy performed in Sir Charles Gairdner Hospital from 1992 to 2012 were included. Exclusion criteria were co-infection with HBV or HIV, previous liver transplantation, hemochromatosis, α1-antitrypsin deficiency, Wilson disease and autoimmune liver diseases. 153 patients who had successful antiviral treatment for HCV were excluded. Patients were followed from the time of baseline liver biopsy until each clinical outcome (death, liver decompensation or HCC) or the end of study. The study was approved by the Sir Charles Gairdner Hospital Human Research Ethics Committee and the Department of Health WA Human Research Ethics Committee.

3.3.2. Liver fibrosis evaluation

Liver biopsy reports were obtained from the database of the Department of Hepatology, Sir Charles Gairdner Hospital. The severity of liver fibrosis was reported by description before 1999, by the Ishak system from 2000 to 2002 and by the Metavir system after 2002. Ishak stage was converted to Metavir stage according to the following scheme: Ishak 0 to F0, Ishak 1, 2 to F1, Ishak 3 to F2, Ishak 4 to F3, Ishak 5, 6 to F4. Biopsy reports before 1999 were reviewed by a specialist liver pathologist and were staged using the Metavir system.
3.3.3. Outcome measurement

The Western Australia Data Linkage Unit linked 1033 patients who had a liver biopsy at Sir Charles Gairdner Hospital to the state-wide hospital morbidity data system, mortality records and cancer register database. Data was extracted for each patient and included date and diagnosis of each hospital admission, date and cause of death and date of HCC diagnosis. The hospital admission diagnosis and the cause of death were recorded using ICD 9 (before 1997) and ICD 10 (after 1997) classification codes.

The primary outcome was death or liver transplantation. The cause of death was categorised into liver related death, behaviour related death and other causes. Liver related death was defined as: death directly attributed to liver failure, variceal bleeding, hepatorenal syndrome or HCC; death in which liver disease was a major contributing factor; or liver transplantation. Behaviour related death was defined as death due to drug or alcohol use, drug overdose, drug or alcohol dependence and intentional self-harm. The secondary outcomes were the first episode of liver decompensation or the diagnosis of HCC. Liver decompensation was defined as ascites, hepatic encephalopathy, variceal bleeding, hepatorenal syndrome, jaundice and spontaneous bacterial peritonitis (R18, K72, I85.0, I98.3, K76.7, K65.9, R17 in ICD code 10 and 789.5, 070.44, 070.41, 456.0, 572.4, 567.23, 782.4 in ICD code 9). Patients with no evidence of decompensation one year prior to biopsy were considered compensated at baseline.
3.3.4. Statistical methods

Correlations were determined using the Spearman correlation coefficient. Survival was assessed using Kaplan Meier curves and log rank test. Cox regression was used to identify predictors of survival. The proportional-hazards assumption was tested based on Schoenfeld residuals. The incidence rate was calculated as the number of events divided by at-risk person time. Two sided p values of <0.05 were considered significant.
3.4. RESULTS

Of the 880 CHC patients who were included in the analysis 833 were compensated at baseline (figure 3.1, table 3.1). 68% were male and the mean age was 40 years. Follow-up totalled 9386 person-years with a mean of 11 years (range 1-20). 173 of patients had F0 (no fibrosis), 383 had F1, 124 had F2, 80 had F3 and 73 had F4 fibrosis. Fibrosis stage was significantly associated with age (r=0.38, p<0.001). During follow-up 122 died or underwent liver transplantation, 28 developed HCC and 58 developed liver decompensation. The mean age at time of death or liver transplantation was 52 years (range: 26-87), at time of HCC diagnosis was 61 years (range: 47-80) and at time of hepatic decompensation was 57 years (range: 39-84). At baseline 15 patients had hepatic decompensation, 20 had HCC, 12 had both decompensation and HCC and these patients were followed separately (figure 3.1).

Behaviour related death accounted for 20% of deaths. Younger age but not fibrosis stage or gender was significantly associated with higher risk of behaviour related death with hazard ratio (HR) of 0.91 (95% CI, 0.86-0.96). Liver related death accounted for 39% of deaths and 31% of these were due to HCC. Age but not gender was significantly associated with liver related survival (HR 1.09: 95% CI, 1.06-1.11). There was no significant difference in liver related survival between those with F0, F1 or F2 fibrosis (p=0.163) with an 18 year survival probability of 99% (95% CI, 96%-100%), 96% (95% CI, 91%-98%) and 94% (95% CI, 83%-98%) respectively (figure 3.2, table 3.3). The liver related survival probability for F3 patients was 96% (95% CI, 88%-99%) at 13 years but decreased to 77% (95% CI, 51%-91%) at 18 years. Age adjusted HR
of liver related death for F3 compared to F0-F2 was 4.24 (table 3.2), with no significant difference in the first 13yr follow up. Those with F4 had significantly worse liver related survival than F3 (p<0.001) with a five, ten and 15 year liver related survival probability of 83% (95% CI, 72%-90%), 57% (95% CI, 43%-70%) and 40% (95% CI, 23%-57%) respectively and the incidence rate of liver related death was 0.051 case/person-year.

Patients with F0, F1 and F2 had excellent HCC free survival with 15 year survival probability of 100%, 99% (95% CI, 98%-100%) and 98% (95%CI, 92%-100%) respectively (figure 3.3, table 3.4). The 15 year decompensation free survival for F0, F1 and F2 patients was 100%, 96% (95%CI, 93%-98%) and 94% (95%CI, 86%-97%) respectively (figure 3.3, table 3.5). F3 patients compared with those with F0-F2 had a significantly reduced decompensation free survival (p=0.009) and HCC free survival (p<0.001). CHC patients with F3 fibrosis had a seven year decompensation free survival of 97% (95% CI, 89%-99%) and seven year HCC free survival of 98% (95% CI, 90%-100%) but this decreased to 86% (95% CI, 70%-94%) and 78% (95% CI, 52%-91%) respectively at 18 years. The age adjusted HR for liver complications (HCC or liver decompensation) free survival for F3 compared to F0-F2 was 3.32 (95% CI, 1.59-6.89), however there was no significant difference during the first seven years follow up (table 3.2). F4 patients had a five and ten year decompensation free survival of 76% (95% CI, 0.63-0.84) and 58% (95% CI, 44%-70%) respectively and HCC free survival of 94% (95% CI, 84%-98%) and 78% (95% CI, 62%-88%) respectively. The age adjusted HR for liver complication free survival for F4 patients compared to those with F0-F2 fibrosis was 17.2 (95% CI, 9.34-31.5). In F4 patients the incidence rate for liver decompensation was 0.068
cases/person-year, for HCC was 0.028 cases/person-year and for liver complications was 0.083 case/person-year.

85 patients with liver decompensation and 60 patients with HCC were identified either at baseline or during follow up and were followed for a mean of two years. 55 decompensated patients and 30 HCC patients had a liver related death with an incidence rate of 0.27 cases/person-year and 0.23 cases/person-year respectively. The one year survival probability for liver related death in decompensated patients was 52% similar to 63% in HCC patients (figure 3.4).
Figure 3.1. Flow chart of patient inclusion and follow up.
Figure 3.2. Liver related survival for compensated patients with each Metavir stage.
Figure 3.3. Clinical outcomes for compensated patients with each Metavir stage. (A) HCC-free survival for each Metavir stage and (B) decompensation-free survival for each Metavir stage.

Figure 3.4. Liver related survival for patients with liver complication. (A) Liver related survival after diagnosis of hepatocellular carcinoma (HCC) and (B) liver related survival after decompensation.
### Table 3.1. Patient characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Baseline fibrosis status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F0</td>
</tr>
<tr>
<td>Number of patients</td>
<td>173 (21)</td>
</tr>
<tr>
<td>Age</td>
<td>36 (8)</td>
</tr>
<tr>
<td>Male</td>
<td>110 (64)</td>
</tr>
<tr>
<td>Female</td>
<td>63 (36)</td>
</tr>
<tr>
<td>Follow up years</td>
<td>12.0 (4.6)</td>
</tr>
<tr>
<td>HCC</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Decompensation</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Total deaths</td>
<td>11 (6.4)</td>
</tr>
<tr>
<td>a. Liver related</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>b. behaviour related</td>
<td>5 (2.9)</td>
</tr>
<tr>
<td>c. other causes</td>
<td>5 (2.9)</td>
</tr>
</tbody>
</table>

**Note:** Continuous variables were expressed as mean (SD), categorical variables were expressed as count (%).

### Table 3.2. Hazard ratio for liver complication and liver related death

<table>
<thead>
<tr>
<th>Metavir stage</th>
<th>Adjust HR*, 95% CI, p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Liver complications</td>
</tr>
<tr>
<td></td>
<td>First 7yrs</td>
</tr>
<tr>
<td>F0-F2</td>
<td>reference</td>
</tr>
<tr>
<td>F3</td>
<td>1.91 (0.51-7.14)</td>
</tr>
<tr>
<td></td>
<td>P=0.33</td>
</tr>
<tr>
<td>F4</td>
<td>20.1 (8.84-45.5)</td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
</tr>
</tbody>
</table>

*Hazard ratio was adjusted for age.
Table 3.3. Liver related survival probability for each Metavir stage

<table>
<thead>
<tr>
<th>Metavir</th>
<th>Accumulative survival probability, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3yr</td>
</tr>
<tr>
<td>F0</td>
<td>0.99</td>
</tr>
<tr>
<td>F1</td>
<td>1.00</td>
</tr>
<tr>
<td>F2</td>
<td>0.99</td>
</tr>
<tr>
<td>F3</td>
<td>0.98</td>
</tr>
<tr>
<td>F4</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table 3.4. HCC free survival probability for each Metavir stage

<table>
<thead>
<tr>
<th>Metavir</th>
<th>Accumulative survival probability, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3yr</td>
</tr>
<tr>
<td>F0</td>
<td>1.00</td>
</tr>
<tr>
<td>F1</td>
<td>1.00</td>
</tr>
<tr>
<td>F2</td>
<td>1.00</td>
</tr>
<tr>
<td>F3</td>
<td>1.00</td>
</tr>
<tr>
<td>F4</td>
<td>0.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metavir</th>
<th>Accumulative survival probability, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3yr</td>
</tr>
<tr>
<td>F0</td>
<td>1·00</td>
</tr>
<tr>
<td>F1</td>
<td>1·00 (0·98-1·00)</td>
</tr>
<tr>
<td>F2</td>
<td>0·98 (0·93-1·00)</td>
</tr>
<tr>
<td>F3</td>
<td>0·99 (0·91-1·00)</td>
</tr>
<tr>
<td>F4</td>
<td>0·84 (0·73-0·91)</td>
</tr>
</tbody>
</table>
3.5. DISCUSSION

This study has comprehensively evaluated the long term clinical sequelae of CHC patients stratified by all liver fibrosis stages (F0-F4). The strengths of this study were the large well defined cohort that included patients with all stages of liver fibrosis, particularly those with none or minimal fibrosis. The long follow up for up to 20 years and the availability of detailed outcomes for each patient has allowed a more precise analysis of the disease burden for CHC patients.

Two previous studies have reported mortality rates related to fibrosis stage for CHC patients. One studied 1050 patients, 93% of whom had Ishak stage 3 to 6, for a maximum of six years and found a significant increase in liver complications and mortality in those with cirrhosis compared to those without (214). The other study included 131 patients with Ishak stage 4 to 6 with a median follow up of 51 months and found no influence of fibrosis stage on liver related survival (215). Others tested the ability of serum fibrosis markers and liver biopsy stage to predict outcomes for hepatitis C patients but these were limited by inclusion of non-CHC patients or a limited 5 year follow up time (211-213).

This study found that CHC patients with no fibrosis, F1 or F2 fibrosis rarely developed progressive disease and they had a minimal risk of developing liver related morbidity and mortality during the following 15 years. The 15 year accumulative HCC free survival, decompensation free survival and liver related survival probability for F2 patients was 98%, 94% and 94% respectively. The implication of these findings is that the use of expensive anti-viral treatments that have significant and severe side-effects may not be justified in those
patients with F0, F1 or F2 fibrosis. Present international guidelines vary in their recommendations regarding HCV treatment based on fibrosis stage. AASLD and EASL guidelines recommend that all patients with F2 fibrosis be strongly considered for or be treated with anti-viral medications, whilst APASL guidelines recommend that genotype 3 patients with any stage of fibrosis be treated and those with genotype 1 be treated if they have moderate or severe fibrosis (233-235). The Australian Pharmaceutical Benefits Scheme guidelines allow federally funded subsidised treatment for CHC patients with any fibrosis stage (236). Clearly these guidelines may need to be reassessed in light of the findings of this present study.

Another important finding was the delayed but significant increase in liver complications in those CHC patients with F3 fibrosis compared with F0, F1 or F2. The increased rate of liver decompensation and HCC development occurred after a delay of seven years follow up and the increased rate of liver related death occurred after 13 years follow up. Previous studies have found that fibrosis progression in CHC does not proceed in a linear manner and that the time needed to progress to cirrhosis from each fibrosis stage was uncertain (19, 41). Based on the findings in this study that liver complications were minimal during 15 years of follow up for F2 patients and occurred after a delay of seven years follow up for F3 patients, it could be assumed that it would take about seven years for F3 patients and more than 15 years for F2 patients to progress to cirrhosis. The use of a repeat liver biopsy or a non-invasive marker of liver fibrosis to monitor for fibrosis progression would be useful in this situation. Surveillance for HCC and oesophageal varicies may also need to be extended to those CHC patients with F3 after a follow up period of seven years.
The natural history of compensated cirrhotic (F4) CHC patients has been more clearly defined. A recent review summarised 13 studies and reported the range of incidence rates (cases/person year) for the development of HCC was 0.018-0.071, for liver complications was 0.028-0.117 and for death or liver transplantation was 0.027-0.067 (47). The results from the Western Australian CHC population in this study fall within these ranges with the corresponding incident rates of 0.028, 0.083 and 0.051 respectively. As expected CHC patients with decompensated liver disease or HCC had significantly worse outcomes than compensated cirrhotic patients. The one year liver related survival probability for decompensated patients was 52% and for HCC was 63% consistent with prior studies (51, 166, 167).

Due to the retrospective nature of the study design, detailed information regarding alcohol consumption and the number of patients who had been unsuccessfully treated with HCV antiviral therapy was not available for analysis. The low percentage of patients with successful treatment was probably due to the high liver biopsy rate and relatively low sustained viral response rate before 2006. Secondly, patient outcome information was extracted from a validated state-wide hospital morbidity, mortality and cancer register database using ICD codes. Data linkage accuracy has been shown to be excellent in previous studies, however there remains the small possibility of miscoding or missing outcomes.

This study has provided new evidence that CHC patients with F0, F1 or F2 fibrosis have a benign course with infrequent episodes of liver related morbidity and mortality after 15 years follow up. Moreover, those CHC patients with F3
fibrosis had a similar low rate of liver complications during the first seven years of follow up after which the rate of complications significantly increased. This period probably represents the time required for those with F3 fibrosis to progress to cirrhosis. This information has significant implications regarding the important public health issues surrounding CHC and the potential cost benefits of antiviral treatment.
CHAPTER 4

IMAGE ANALYSIS OF LIVER COLLAGEN USING SIRIUS RED IS MORE ACCURATE AND CORRELATES BETTER WITH SERUM FIBROSIS MARKERS THAN TRICROME

4.1. ABSTRACT

4.1.1. Background

Collagen proportional area (CPA) determined by quantitative digital image analysis better quantifies liver fibrosis than histological stage, however its clinical use has been limited by non-standardized methods.

4.1.2. Aims

This study aimed to compare CPA obtained using different staining methods, magnifications and biopsy sizes.

4.1.3. Methods

249 patients with chronic hepatitis C (CHC) who had a liver biopsy and serum fibrosis markers performed were included. CPA was measured either using a sirius red (CPAs) or a trichrome (CPAt) stain.
4.1.4. Results

CPAs measured at 20X and 40X magnifications generated similar outcomes with inter-class correlation (ICC) coefficient of 0.98. Compared to trichrome, sirius red staining had much less variation with an ICC coefficient of 0.99 for slides stained in the same batch and 0.92 in different batches. Mean CPAs was higher than mean CPAt by 3.53%, P<0.001. Morphological analysis found that sirius red detected delicate fibrous septa and spurs better than trichrome. Both CPAs and CPAt correlated well with Metavir stage while CPAs had better ability to detect cirrhosis with the area under ROC curve of 0.95. Overall CPA had superior correlation with serum markers of fibrosis in Metavir F2-F4 than that in F0-F1 and CPAs correlated better with serum fibrosis markers than CPAt in Metavir F0-F1. Multivariate analysis found that HA, α2-macroglobulin, platelet count and albumin were independently correlated with CPAs and only HA was independently correlated with CPAt.

4.1.5. Conclusions

Sirius red staining for CPA determination was more accurate and reliable for quantifying hepatic collagen compared with trichrome staining.
4.2. INTRODUCTION

Histological staging systems such as Metavir have long been the reference standard to assess liver fibrosis. However, traditional staging systems have well documented limitations including inter and intra-observer variability and a small number of categories (237). There is increasing evidence that important additional information is hidden within each histological stage, especially in advanced fibrosis and cirrhosis. In the setting of cirrhosis, additional histological features including nodular size and fibrous septa thickness were showed to independently predict the presence of clinically significant portal hypertension (172). Understanding the potential value of a range of histological features in addition to Metavir fibrosis stage is essential in order to better predict outcomes and individualize therapy (164).

Quantitative digital image analysis (DIA) is a newly developed method used to assess liver fibrosis. DIA segments digital images of liver biopsies and accurately measures the area of collagen and the area of remaining liver tissue and produces the proportion of the area of the biopsy occupied by collagen [collagen proportional area (CPA)] (73). Unlike traditional histological staging systems which mainly depend on the degree of architectural distortion, DIA measures the amount of fibrosis as a continuous variable. Initial studies found that CPA was well correlated with histological stage (75, 76, 238, 239). Furthermore, as a computer-based technology, CPA was more reproducible due to minimised inter and intra-observer variability of liver biopsy interpretation.

Up to the present, a number of different technologies and methods have been used to obtain the CPA. Masson’s trichrome is a routine staining method used
for liver fibrosis and this was used in studies by O’Brien et al (77), Lazzarini et al (85) and Sethasineet al (86). Other studies have used sirius red staining because of its specific binding to collagen (68, 75, 87-89). Pixel counting technology is potentially a highly accurate technology that calculates collagen area, however, its reliability could be influenced by threshold settings, magnification and image resolution. A wide range of magnifications have been used and range from 1X by Calvaruso et al (76) to 100X by Bedossa et al (78). The minimum biopsy size for CPA measurement was also unclear. These variables limit direct comparison between studies and clinical interpretation of CPA values. Therefore there is a need to standardise each stage of CPA analysis to minimise variability and thus broaden the clinical use of CPA.

Serum markers models such as Hepascore (122), APRI (120), Lok index (129) and FIB-4 (123) are increasingly used clinically as non-invasive methods to detect the severity of liver fibrosis. Similar to CPA, these models measure liver fibrosis in a continuous manner. There is some evidence that CPA is correlated with serum fibrosis and biochemical markers although there are conflicting results (75, 133). The differences between studies may be due to different methods used to measure CPA and different patient populations.

The aim of this study was to compare CPA obtained using different staining methods, magnifications and biopsy sizes and to determine the optimum method of CPA measurement. The correlation of CPA with Metavir stage, serum fibrosis markers and models was also evaluated.
4.3. METHODS

4.3.1. Patients

249 chronic hepatitis C patients from the liver clinics at Sir Charles Gairdner Hospital who had a liver biopsy and a Hepascore (122) performed within 6 months of biopsy from January 1997 until January 2012 were retrospectively included. Chronic HCV infection was defined as the presence of hepatitis C RNA on two separate occasions more than 6 months apart. Exclusion criteria included co-infection with HBV and HIV; other liver diseases including hemochromatosis, α1-antitrypsin deficiency, Wilson’s disease and autoimmune liver diseases; successful anti-viral treatment before biopsy; liver transplantation. The study was approved by the Sir Charles Gairdner Hospital Human Research Ethics Committee.

4.3.2. Liver biopsy staining

Biopsy samples were formalin fixed, paraffin embedded and routinely stained with Masson’s trichrome. A further tissue section was cut and stained with sirius red. To test the reproducibility of the staining methods, two consecutively cut sections from 10 randomly chosen tissue blocks were stained in the same batch and another two consecutively cut sections were stained in different batches.

4.3.3. Quantitative image analysis

Liver biopsies were stained with trichrome and sirius red and scanned using the Aperio Scanscope XT Digital Slide Scanner at 40X magnification. (1.59X10^7 Pixels = 1 mm^2). The image was viewed using Aperio ImageScope software
version 10.0. The liver capsule and large portal tracts with a diameter greater than 500 um were excluded as these did not represent disease related collagen (172). The optimum threshold for positive pixels that corresponded to areas of sirius red or trichrome staining was determined using the software by changing the hue value and colour saturation using the original image for comparison. A binary image was produced and the CPA was expressed as a percentage of positive pixels to total pixels. The liver biopsy area measured was recorded. CPA reproducibility was tested by repeating the analysis four times on one slide each with high, intermediate and low CPA values. To test the influence of magnification, 20 randomly selected sirius red stained biopsies were scanned and analysed at both 20X and 40X magnifications. Finally variability due to biopsy size was studied using two large surgical specimens of liver with chronic hepatitis C (Metavir F4). Sections were stained with sirius red, scanned at 40X magnification and a test area 5mm wide and 20mm long was chosen for analysis. This area was sub-divided equally into five 1mm wide fields and each field was sequentially sub-divided into ten lengths of 2mm, four lengths of 5mm, two lengths of 10mm or one length of 20mm. CPA was measured for the test area and each of the four different sized biopsy areas.

### 4.3.4. Clinical data

Clinical and laboratory data of patients including age, gender, Hepascore, Metavir stage, hyaluronic acid (HA), bilirubin, gamma-glutamyl transpeptidase (GGT), α2-macroglobulin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelet count, prothrombin time, international normalized ratio (INR), alkaline phosphatase (ALP) and albumin, were obtained
from the medical database of the Department of Gastroenterology and Hepatology, Sir Charles Gairdner Hospital. APRI results were calculated as aspartate aminotransferase (AST) (U/L)/upper normal X 100/platelet count (10⁹/L) (120). Lok index results were calculated as -5.56-0.0089Xplatelet (10³/mm³)+1.26XAST/ALT+5.27XINR (129). FIB-4 results were calculated as age (years) X AST [U/L]/((platelets [10⁹/L] X (ALT [U/L])¹/²) (123).

4.3.5. Statistical analysis

The paired sample t-test was used to compare the CPA values obtained from sirius red (CPAs) and trichrome staining (CPAt). The Mann-Whitney U test was used to compare medians. Inter-class correlation (ICC) was used to evaluate test variability. Coefficient of variance (CV) was calculated as the ratio of standard deviation against the mean value. Correlations between CPA and Metavir stage were determined by Spearman correlation coefficients. Correlations between CPA and serum fibrosis markers were determined by Pearson correlation coefficients.
4.4. RESULTS

249 patients were initially included: 187 males (75%) and 62 females (25%); mean age was 42 years old; 19% of patients had Metavir F0, 36% had F1, 22% had F2, 14% had F3 and 9% had F4 (table 4.1).

Two surgical specimens were used to evaluate the variability of CPAs related to different biopsy size. The CPAs was measured for the 100 mm$^2$ test area of each specimen and was 25.5% for patient 1 and 11.9% for patient 2 (figure 4.1A and 4.1B). The CPAs for the four different sized biopsy areas were then measured and compared with these values. The variability of CPAs progressively decreased with larger measurement areas, e.g. CV of 52.6% for 2 mm$^2$, 37.1% for 5 mm$^2$, 34.9% for 10 mm$^2$ and 16.5% for 20 mm$^2$ in patient 1. A similar trend was observed in patient 2. The results from both patients were combined to determine the relative variance for each biopsy area. The relative variance was defined as the difference between the sub-area CPAs and the test CPAs divided by the test CPAs (figure 4.1C). Almost all biopsies with a size $\geq$ 5mm$^2$ had a relative variance less than 50%. There was no obvious difference in relative variance between a biopsy area of 5mm$^2$ and 10mm$^2$.

The effect of magnification on CPA was tested and the ICC coefficient of CPAs measured at 20X and 40X magnifications was 0.98. The resolution of the image scanned at 20X magnification was $3.94 \times 10^6$ pixels=1 mm$^2$ and showed excellent accuracy in measuring CPAs. The time needed for an image scan and analysis was doubled for 40X magnification ie from 3 to 6 mins for each biopsy with 10mm length.
As a result of the studies documented above, 48 patients with a liver biopsy size < 5mm$^2$ were excluded from the final analysis that compared sirius red staining with trichrome. The 201 patients in the final analysis had a mean biopsy area of 10.41mm$^2$ for CPAs measurement and 12.16 mm$^2$ for CPAt measurement (table 4.1). The median CPAs was 6.18% in biopsies sized from 5mm$^2$ to 10mm$^2$ and 6.0% in biopsies sized larger than 10mm$^2$. Paired comparison of CPAs and CPAt found that the mean CPAs was higher than mean CPAt by 3.53% (95% CI: 2.87-4.20%, P<0.001). Stratified analysis of results found a significant difference between CPAs and CPAt only in patients with Metavir F0 to F3 stage of fibrosis. The paired values of CPAs and CPAt for each patient using CPAs as the reference value are shown in figure 4.2. CPAt was generally lower than CPAs for each patient and the difference increased in patients with higher CPAs.

Collagen that was present around portal tracts was successfully stained by both methods, however sirius red was better able to detect fibrous septa or spurs that extended into the hepatic lobule (figure 4.3). The difficulties in setting the threshold for positive pixels in trichrome stained biopsies also lead to further variation. Viewing trichrome stained sections, collagen was coloured blue and nuclei were coloured either black, brown or blue. Therefore in some biopsies there was an obvious overlap of colours staining collagen and nuclei and this caused either an overestimation or underestimation of collagen quantification. In contrast, sirius red stained sections used only one dye that bound specifically to collagen and this resulted in a higher contrast that allowed improved accuracy in setting thresholds for positive pixels.
The variability of the staining and scanning techniques was tested and the ICC coefficient for variation of sirius red staining was 0.99 for sections stained in the same batch and 0.92 for sections stained in different batches. This was significantly better than the variability of trichrome staining where the ICC coefficient was 0.51 and 0.60 respectively. The reproducibility of image analysis was excellent for both methods where the ICC coefficient was 0.990 for sirius red stained biopsies and 0.999 for trichrome stained biopsies.

Both CPAs and CPAt correlated well with Metavir stage (CPAs: r=0.63, p<0.001, CPAt: r=0.55, p<0.001). As expected there was a large range of CPAs and CPAt values within Metavir F3 and F4 stages and this resulted in considerable overlap between CPA values in F3 and F4 biopsies (figure 4.4). The median CPAs was significantly different comparing Metavir F1 and F2 (p<0.001); F2 and F3 (p=0.011); F3 and F4 (p<0.001). Whereas the median CPAt was only significantly different comparing Metavir F1 and F2 (p=0.004); F2 and F3 (p=0.001). The area under the ROC curve (AUROC) for CPAs to detect cirrhosis (Metavir F4) was 0.95 (95% CI, 0.90-0.99). A CPAs value of 12.8% resulted in a sensitivity of 81.3% and specificity of 90.3% for detecting F4 (figure 4.4). CPAt was less predictive for cirrhosis with the AUROC of 0.86 (95% CI, 0.78-0.95).

Eleven individual serum markers (bilirubin, GGT, hyaluronic acid, α2-macroglobulin, ALT, AST, platelet count, prothrombin time, INR, ALP and albumin) were analysed for correlation with CPA. CPA, GGT and HA were log transformed to attain a normal distribution Using univariate analysis seven markers showed significant correlation with both CPAs and CPAt (table 4.2).
Stratification by Metavir stage found that $\alpha$2-macroglobulin, HA and platelet count was significantly correlated with CPAs in both F0-F1 and F2-F4 (table 4.2). GGT, albumin, ALP and INR were only significantly correlated with CPAs in Metavir F2-F4. No serum marker was found to be significantly correlated with CPAt in Metavir F0-F1 while eight serum markers had significant correlation with CPAt in Metavir F2-F4. Multivariate analysis found that HA, $\alpha$2-macroglobulin, platelet count and albumin were independently predictive for CPAs and only HA was independently correlated with CPAt (table 4.3). The four serum fibrosis models (Hepascore, APRI, FIB-4 and Lok index) were all significantly correlated with both CPAs and CPAt and Hepascore had the highest correlation coefficient of 0.56 (table 4.4). CPAs had higher correlation coefficients with the serum fibrosis models than CPAt. Hepascore and Lok index were significantly correlated with CPAs in both F0-1 and F2-4 groups while FIB-4 and APRI were only significantly correlated with CPAs in F2-F4.
Figure 4.1. CPAs variation associated with size of measurement area using biopsies from two patients. (A) Patient 1: CPAs for whole measurement area was 25.5%. (B) Patient 2: CPAs for whole measurement area was 11.9%. (C) Relative variance for each measurement area combining both biopsy specimens. The relative variance was defined as the difference between the sub-area CPAs and the test CPAs divided by the test CPAs.
Figure 4.2. CPAt and CPAs for each patient. Patients were sorted according to CPAs.
<table>
<thead>
<tr>
<th>Metavir</th>
<th>Sirius Red</th>
<th>Trichrome</th>
<th>CPA values</th>
</tr>
</thead>
</table>
| F0      | ![Image](image1) | ![Image](image2) | CPAs=4.55%  
CPAt=0.3% |
| F1      | ![Image](image3) | ![Image](image4) | CPAs=4.09%  
CPAt=1.31% |
| F2      | ![Image](image5) | ![Image](image6) | CPAs=8.34%  
CPAt=2.52% |
| F3      | ![Image](image7) | ![Image](image8) | CPAs=13.59%  
CPAt=7.17% |
| F4      | ![Image](image9) | ![Image](image10) | CPAs=22.08%  
CPAt=4.73% |

Figure 4.3. Liver biopsies stained using sirius red and trichrome for each Metavir stage.
Figure 4.4. CPAs and CPAt grouped according to Metavir stage in 201 CHC patients. A CPAs value of 12.8% resulted in a sensitivity of 81.3% and specificity of 90.3% for detecting F4. ○: outliers.
Table 4.1. Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Whole cohort (n=249)</th>
<th>After exclusion (n=201)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, n (%)</td>
<td>62 (25)</td>
<td>47 (23)</td>
</tr>
<tr>
<td>age, years</td>
<td>42 (10)</td>
<td>42 (10)</td>
</tr>
<tr>
<td>CPAs, (%)</td>
<td>-</td>
<td>6.1 (1.8-32.7)</td>
</tr>
<tr>
<td>CPAt, (%)</td>
<td>-</td>
<td>2.5 (0.03-33.9)</td>
</tr>
<tr>
<td>GGT (U/L)</td>
<td>58 (8-662)</td>
<td>54 (8-662)</td>
</tr>
<tr>
<td>HA (ug/L)</td>
<td>29 (1-697)</td>
<td>28 (1-697)</td>
</tr>
<tr>
<td>platelet count (10⁹/L)</td>
<td>218 (79)</td>
<td>218 (67)</td>
</tr>
<tr>
<td>α2-macroglobulin (g/L)</td>
<td>2.9 (1.1)</td>
<td>2.9 (1.2)</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>127 (97)</td>
<td>126 (94)</td>
</tr>
<tr>
<td>prothrombin time (secs)</td>
<td>9.0 (1.6)</td>
<td>8.9 (1.7)</td>
</tr>
<tr>
<td>Albumin (g/L)</td>
<td>43 (4)</td>
<td>43 (3)</td>
</tr>
<tr>
<td>ALP (U/L)</td>
<td>86 (59)</td>
<td>87 (63)</td>
</tr>
<tr>
<td>INR</td>
<td>1.0 (0.6)</td>
<td>1.0 (0.7)</td>
</tr>
<tr>
<td>Bilirubin (umol/L)</td>
<td>11 (7)</td>
<td>11 (5)</td>
</tr>
<tr>
<td>APRI</td>
<td>1.13 (1.27)</td>
<td>1.11 (1.32)</td>
</tr>
<tr>
<td>Lok index</td>
<td>-1.26 (1.09)</td>
<td>-1.37 (0.90)</td>
</tr>
<tr>
<td>FIB-4</td>
<td>2.0 (2.04)</td>
<td>1.78 (1.60)</td>
</tr>
<tr>
<td>Hepascore</td>
<td>0.49 (0.32)</td>
<td>0.49 (0.32)</td>
</tr>
<tr>
<td>Metavir F0/ F1/F2/ F3/F4 (%)</td>
<td>19/36/22/14/9</td>
<td>17/39/22/14/8</td>
</tr>
</tbody>
</table>

Note: age, platelet count, α2-macroglobulin, ALT, prothrombin time, albumin, ALP, INR, Bilirubin, Hepascore, APRI, Lok index and FIB-4 were expressed as mean (Standard deviation). CPAs, CPAt, HA and GGT were expressed as median (range). After exclusion, 97 APRI and FIB-4 results were available and 94 Lok index results were available. * excluded those with biopsies < 5mm².
Table 4.2. Correlation between serum markers and CPA

<table>
<thead>
<tr>
<th>serum marker</th>
<th>Pearson correlation coefficient (P value)</th>
<th>Total patients (n=201)</th>
<th>F0-F1 (n=112)</th>
<th>F2-F4 (n=89)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IgCPAs</td>
<td>IgCPAt</td>
<td>IgCPAs</td>
<td>IgCPAt</td>
</tr>
<tr>
<td>α2-macroglobulin</td>
<td>0.51 (p &lt; 0.001)</td>
<td>0.30 (p &lt; 0.001)</td>
<td>0.24 (P = 0.011)</td>
<td>0.04 (NS)</td>
</tr>
<tr>
<td>IgHA</td>
<td>0.52 (p &lt; 0.001)</td>
<td>0.40 (p &lt; 0.001)</td>
<td>0.24 (P = 0.011)</td>
<td>0.14 (NS)</td>
</tr>
<tr>
<td>platelet count</td>
<td>-0.49 (p &lt; 0.001)</td>
<td>-0.26 (P = 0.002)</td>
<td>-0.3 (P = 0.007)</td>
<td>0.05 (NS)</td>
</tr>
<tr>
<td>IgGGT</td>
<td>0.24 (P = 0.001)</td>
<td>0.21 (P = 0.003)</td>
<td>0.05 (NS)</td>
<td>-0.01 (NS)</td>
</tr>
<tr>
<td>ALT</td>
<td>0.15 (P = 0.037)</td>
<td>0.16 (P = 0.024)</td>
<td>-0.01 (NS)</td>
<td>-0.01 (NS)</td>
</tr>
<tr>
<td>AST</td>
<td>0.12 (NS)</td>
<td>0.18 (P = 0.025)</td>
<td>-0.11 (NS)</td>
<td>-0.01 (NS)</td>
</tr>
<tr>
<td>prothrombin time</td>
<td>0.32 (P = 0.001)</td>
<td>0.30 (P = 0.002)</td>
<td>0.10 (NS)</td>
<td>-0.07 (NS)</td>
</tr>
<tr>
<td>albumin</td>
<td>-0.23 (P = 0.002)</td>
<td>-0.17 (P = 0.025)</td>
<td>0.06 (NS)</td>
<td>-0.11 (NS)</td>
</tr>
<tr>
<td>ALP</td>
<td>0.04 (NS)</td>
<td>0.11 (NS)</td>
<td>-0.04 (NS)</td>
<td>0.08 (NS)</td>
</tr>
<tr>
<td>INR</td>
<td>0.06 (NS)</td>
<td>0.04 (NS)</td>
<td>0.14 (NS)</td>
<td>0.07 (NS)</td>
</tr>
<tr>
<td>bilirubin</td>
<td>0.05 (NS)</td>
<td>0.03 (NS)</td>
<td>-0.12 (NS)</td>
<td>-0.08 (NS)</td>
</tr>
</tbody>
</table>

Note: CPAs, CPAt, GGT and HA were log transformed. NS: not significant.
Table 4.3. Multivariate analysis of serum markers and CPA

<table>
<thead>
<tr>
<th>Serum marker</th>
<th>Multivariate coefficient (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IgCPAs</td>
</tr>
<tr>
<td>IgHA</td>
<td>0.127 (p=0.014)</td>
</tr>
<tr>
<td>α2-macroglobulin</td>
<td>0.037 (p=0.040)</td>
</tr>
<tr>
<td>platelet count</td>
<td>-0.001 (p=0.001)</td>
</tr>
<tr>
<td>albumin</td>
<td>-0.012 (p=0.027)</td>
</tr>
</tbody>
</table>

**Note:** CPAs, CPAt and HA were log transformed.

Table 4.4. Correlation between serum models and CPA

<table>
<thead>
<tr>
<th>serum model</th>
<th>Pearson correlation coefficient (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total patients (n=201) F0-F1 (n=112) F2-F4 (n=89)</td>
</tr>
<tr>
<td></td>
<td>IgCPAs IgCPAt IgCPAs IgCPAt IgCPAs IgCPAt</td>
</tr>
<tr>
<td>Hepascore</td>
<td>0.56 (p&lt;0.001) 0.39 (p&lt;0.001) 0.30 (P=0.001) 0.02 (NS) 0.40 (p&lt;0.001) 0.35 (P=0.001)</td>
</tr>
<tr>
<td>APRI</td>
<td>0.46 (p&lt;0.001) 0.34 (p&lt;0.001) 0.02 (NS) 0.09 (NS) 0.46 (p&lt;0.001) 0.32 (p&lt;0.001)</td>
</tr>
<tr>
<td>Lok index</td>
<td>0.51 (p&lt;0.001) 0.37 (p&lt;0.001) 0.32 (P=0.038) 0.13 (NS) 0.35 (P=0.011) 0.24 (NS)</td>
</tr>
<tr>
<td>FIB-4</td>
<td>0.55 (p&lt;0.001) 0.41 (p&lt;0.001) 0.29 (NS) 0.16 (NS) 0.49 (p&lt;0.001) 0.36 (P=0.008)</td>
</tr>
</tbody>
</table>

**Note:** CPAs, CPAt were log transformed. NS: not significant.
4.5. DISCUSSION

To date non-standardized methods of liver biopsy staining and measurement of CPA have greatly limited the clinical use of CPA. This is the first study that has systemically evaluated the influence of staining method, scanning technique and biopsy size on the reliability of CPA measurement.

The trichrome stain is used to detect all connective tissue components and not just collagen. The method is based on the sequential application of dyes that results in the replacement of a red dye from less intensely stained components such as collagen and other connective tissues with a poly acid (usually phosphotungstic acid)(240). In contrast the sirius red stain is based on the application of a single dye that has been shown to specifically stain collagen types I, II and III (241) and it was highly sensitive in detecting small amounts of collagen (242). In a paired comparison this present study analysed the CPA obtained with trichrome staining with the CPA obtained using sirius red staining and found that the reproducibility of CPAs was superior to that achieved with CPAt. Moreover the mean CPAs was found to be significantly greater than CPAt and this was consistent with the morphological findings that sirius red staining detected more collagen within the hepatic lobule than did trichrome staining. The higher sensitivity of CPAs not only gave it a significant advantage in quantitative analysis but it was also able to identify subtle changes of liver fibrosis that maybe relevant in assessing either fibrosis progression or regression with treatment. This is important as there is increasing evidence that shows fibrosis regression and improved clinical outcomes occur after successful antiviral treatment for patients with HCV and HBV infection (44, 45, 243-245).
This study also found that the CPAs obtained using either 20X or 40X magnification were similar and the ICC coefficient was 0.98. A previous study conducted by Lazzarini et al found similar CPA results when comparing 40X and 100X magnification however a comparison with 20X magnification was not made (85). The influence of magnification mainly depended on the resolution of the scanner. The resolution of the Aperio Scanscope XT Digital Slide Scanner used in this study was $3.94 \times 10^6$ pixels/mm$^2$ at 20X magnification and this was high enough to guarantee that only one colour was contained in each pixel. A number of different digital image capture technologies and analysis software have been used in previous studies and these have a wide range of image resolution, varying from $2.3 \times 10^5$/mm$^2$ (76) to $1.3 \times 10^6$ per image (85) and image magnifications from 1X (76) to 100X (78). Therefore the accuracy of CPA measurements using these technologies may not be as reliable as that used in this study. Magnification can also affect the threshold setting. Threshold was determined manually by comparing binary images with original images in this and most other studies and higher magnification allowed improved accuracy. It is therefore recommended that a magnification of at least 10X be used to optimise threshold settings.

A critical variable in the assessment of CPA is the minimum biopsy size that can give accurate and reliable results. The CPA coefficient of variance of 37% for a 5 mm$^2$ biopsy area was similar to the CPA coefficient of variance of 35% for a 10mm$^2$ biopsy area and the relative variance was less than 50% for almost all biopsies $\geq$5mm$^2$. Similar median CPAs values among biopsies sized 5mm$^2$ to 10mm$^2$ and biopsies sized larger than 10mm$^2$ also suggested that biopsies sized larger than 5mm$^2$ were acceptable for CPAs measurement. These findings
contrast with the larger variance of CPA found by others where it was suggested that an adequate biopsy length was 40mm (78). The improvement in CPA measurement from smaller biopsy sizes was partly due to the exclusion of large portal tracts and also the use of high resolution scanner. The requirement of a minimum number of portal tracts is less relevant when using CPA analysis as CPA is a continuous variable that does not require morphological information related to portal tracts and indeed large portal tracts are excluded from the analysis. Although a smaller liver biopsy size was acceptable for CPAs measurement, this size was inadequate for other purposes such as Metavir staging and etiological diagnosis.

As expected, CPAs and CPAt were both significantly correlated with Metavir stage. This finding was supported by other studies where the correlation coefficient ranged from 0.48 to 0.84 (75-77, 88, 133, 238). Importantly, this present study found that CPAs was better correlated with Metavir stage than CPAt. The median CPAs was significantly different between each consecutive Metavir stage from F1 to F4. In addition a wide range of CPAs was observed in Metavir F4 and the median CPAs value of 18.3% almost doubled the median CPAs of 9.8% in Metavir F3. This indicated that a significant proportion of cirrhotic patients had an extremely high CPAs and potentially had a greater risk of adverse clinical outcomes.

Eight serum markers (GGT, hyaluronic acid, α2-macroglobulin, ALT, AST, platelet count, prothrombin time, and albumin) were significantly correlated with CPA. Although CPAs was only correlated with seven of these, overall the correlation coefficient was higher than that for CPAt. In particular CPAt was not
significantly correlated with any serum marker in Metavir F0-F1. Multivariate analysis showed that HA, α2-macroglobulin, platelet count and albumin remained significantly correlated with CPAs. Using similar multivariate analysis only HA has previously been shown to be correlated with CPA (75). Another study separated patients into two groups of CPA levels (upper 31%, lower 69%) and found using multivariate analysis that TIMP-1, ALP ratio and platelet significantly correlated with this cut point (133). Importantly, both prior studies were limited by a restricted range of CPA values in that the majority of their patients had advanced fibrosis. The widespread Metavir stages and CPA values in this present study allows better interpretation of the association between serum markers and CPA values in a broad clinical setting.

All four serum models correlated well with CPAs and less well with CPAt. Hepascore had the highest correlation coefficient with CPAs and significantly correlated with CPAs in both groups of Metavir F0-F1 and F2-F4. This suggested that in addition to detecting significant fibrosis or cirrhosis, serum markers and models had the ability to provide extra information within each Metavir stage.

In summary, this study conducted a paired comparison of CPA measurement using different biopsy sizes, magnifications and staining methods. CPA obtained using sirius red had higher accuracy and reproducibility compared to trichrome and better correlated with Metavir stage, serum markers and models. Further investigations will be required to confirm these findings.
CHAPTER 5

IMAGE ANALYSIS OF LIVER BIOPSY SAMPLES
MEASURES FIBROSIS AND PREDICTS CLINICAL OUTCOME

5.1. ABSTRACT

5.1.1. Background and Aims

Histopathological scoring of liver fibrosis mainly measures architectural abnormalities and requires a minimum biopsy size (≥10mm). Liver collagen quantification may allow use of small size biopsies and improve the prediction of clinical outcomes. This study evaluated the ability of the collagen proportional area (CPA) measurement to predict clinical outcomes.

5.1.2. Methods

Clinical outcomes were determined using population based data-linkage for chronic hepatitis C (CHC) patients from 1992-2012. Quantitative digital image analysis of liver biopsies was used for CPA measurement.

5.1.3. Results

533 patients with a biopsy size ≥5 mm were included. Median follow up was 10.5 years. 26 developed hepatocellular carcinoma (HCC), 39 developed liver decompensation and 33 had liver related death. 453 had Metavir F0-F2 and 80
had F3-F4. CPA ranged from 1.3%-44.6%. CPA and Metavir stage were independently associated with liver related death. Metavir stage, CPA stage and age were independently associated with HCC. CPA stage (C1: 0%-5%, C2: 5%-10%, C3: 10%-20%, C4: >20%) stratified risk and a significant difference in outcomes was present between all CPA stages for HCC and between C2-C3 and C3-C4 for decompensation and liver related death. The 15 year composite endpoint-free survival was 97% for C1, 89% for C2, 60% for C3, 7% for C4. C4 had significantly worse survival than ≤C3 (p<0.001) in cirrhotic patients.  

5.1.4. Conclusions  

CPA stage gave additional information regarding risk stratification for adverse clinical outcomes independent of Metavir stage.
5.2. INTRODUCTION

Development of liver fibrosis is a critical feature of progressive chronic hepatitis C (CHC) and its severity is closely associated with adverse clinical outcomes (246). The evaluation of liver fibrosis severity is essential to determine prognosis and guide management of CHC patients. Liver biopsy histological staging systems such as Metavir or Ishak have long been used to assess the severity of liver fibrosis (247) and the risk of liver related morbidity and mortality significantly accelerate after the development of cirrhosis (Metavir F4) (214). However, the correlation between histological stage and clinical outcomes in patients with less severe fibrosis is not clearly determined. Additional histological features not incorporated into the classical fibrosis staging systems may also have important prognostic implications. A study showed that nodule size and fibrous septa thickness were independent predictors for the presence of clinically significant portal hypertension (172).

Quantitative digital image analysis is a newly developed method used to assess liver fibrosis. This technology segments digital images of liver biopsies and accurately measures the area of collagen and the area of remaining liver tissue and calculates the proportion of the biopsy occupied by collagen [collagen proportional area (CPA)] (73). In contrast to histological staging systems CPA is a continuous measure of the amount of liver fibrosis and has minimal inter and intra-observer variability. Non standardized image capture techniques and image analysis methods have been used for CPA measurement and this has limited the use of CPA to a degree. Recently we evaluated and optimized the method for CPA measurement using small liver biopsy samples and achieved a high degree of accuracy and reproducibility (248).
Despite these limitations early studies have shown a significant correlation between CPA and histological stage as well as CPA and serum markers of liver fibrosis (75, 76, 133, 239). CPA also correlated with hepatic venous pressure gradient in CHC patients (76), with decompensation in cirrhotic patients (81), with transplantation free survival in children with biliary atresia (249) and post liver transplantation outcomes for recurrent hepatitis C virus (HCV) (82-84). Therefore CPA has the potential to better predict clinical outcomes for CHC patients than established histological scoring systems and allow improved management of these patients. The aim of this study was to compare the ability of the optimised method of CPA measurement with Metavir stage to predict long term liver related morbidity and mortality in a large well documented cohort of CHC patients with a range of fibrosis stages from none (F0) to cirrhosis (F4).
5.3. METHODS

5.3.1. Patients

All CHC patients who attended the Department of Gastroenterology/Hepatology, Sir Charles Gairdner Hospital and had a liver biopsy from 1992 to 2012 were included. An archived liver tissue block needed to be available for use. Exclusion criteria included co-infection with hepatitis B virus and human immunodeficiency virus; other liver diseases including hemochromatosis, α1-antitrypsin deficiency, Wilson disease and autoimmune liver diseases; and a history of liver transplantation, decompensation or diagnosis of HCC before liver biopsy. Patients who were successfully treated for HCV were also excluded. The study was approved by the Sir Charles Gairdner Hospital Human Research Ethics Committee and the Department of Health Human Research Ethics Committee.

5.3.2. Liver fibrosis evaluation

Liver biopsies were obtained percutaneously and routinely stained using Masson’s trichrome. All biopsy slides were reviewed by an expert liver pathologist (BdB) who was blinded to clinical data. Liver fibrosis was staged using the Metavir staging system, namely: F0, no fibrosis; F1, portal fibrosis without septa; F2, few septa; F3, numerous septa without cirrhosis; F4, cirrhosis (250). The size, fragmentation and number of portal tracts of each biopsy was recorded.
5.3.3. Quantitative image analysis

A new liver section was cut from the stored tissue block and stained using sirius red and this was used for CPA analysis. CPA was measured using the optimised method previously described (248). Sections were scanned using the Aperio Scanscope XT Digital Slide Scanner at 20X magnification ($1.59 \times 10^7$ Pixels = 1 mm$^2$). The image was viewed using Aperio ImageScope software version 10.0. The liver capsule and large portal tracts were excluded as these did not represent disease related collagen (172). The optimum threshold for positive pixels that corresponded to areas of sirius red staining was determined according to hue value and colour saturation using the original image for comparison. A binary image was produced by the software and CPA was expressed as a percentage of positive pixels to total pixels. The CPA measurement was calculated by the software. According to our previous study, a biopsy with a measurement area less than 5 mm$^2$ after exclusion of any large portal tracts or capsule was considered as insufficient and thus excluded from further analysis (248).

5.3.4. Data source

The long term follow up of patients was obtained from the Western Australia Data Linkage Unit. This is a validated population-based data linkage system that links multiple health related datasets including the state cancer register, the state hospital morbidity database and the state mortality records dating back to 1982, 1970 and 1969 respectively (229). The Hospital Morbidity Data System has 100% coverage of data for hospital admissions throughout the state with a record linkage success rate of more than 99% (229). The hospital admission
diagnosis and the cause of death were recorded using ICD 9 (before 1997) and ICD 10 (after 1997) classification codes. Personal identifiers were encrypted and stored separately from the data used for analysis.

5.3.5. Endpoints and statistical analysis

The primary endpoint was liver related death or liver transplantation. Liver related death was defined as death from liver failure, variceal bleeding or hepatocellular carcinoma as well as death in which liver disease was the major contributing factor. The secondary endpoints were the first episode of liver decompensation (ascites, hepatic encephalopathy, variceal bleeding, hepatorenal syndrome, spontaneous bacterial peritonitis) or the development of HCC. Survival duration was determined from the time of liver biopsy to each endpoint. The correlation between CPA and Metavir stage was assessed by the Spearman correlation coefficient. Survival was assessed using Kaplan Meier curves and significance determined by the log rank test and cox regression analysis. Area under receiver operating characteristic curves (AUROC) was calculated for CPA and Metavir stage to predict each endpoint. Two sided p values of <0.05 were considered significant.
5.4. RESULTS

844 CHC patients were initially included. The stored liver biopsy block used for sirius red staining and CPA measurement had been sectioned previously for routine histopathology assessment and the remaining stored tissue was therefore of smaller size and often <5 mm. Therefore, 208 patients were excluded due to insufficient liver biopsy size. 103 patients who had successful antiviral treatment were also excluded from the core analysis but were included in the sensitivity analysis. 533 patients were included in the final analysis with a median follow up time of 10.5 years (range 0.1-20 years) (table 5.1). The mean biopsy length used for Metavir staging was 15mm (SD: 4.4mm) and the mean portal tract number was 9 (SD: 4.0). Less than 1% of biopsies were fragmented. The mean biopsy length used for CPA measurement was 10mm (SD: 3.7mm). The range of CPA values varied from 1.3% to 44.6%. CPA value was well correlated with Metavir stage with a correlation coefficient of 0.7377, P<0.001. There was a wide range of CPA values within each Metavir stage with a median CPA of 3.7% for F0, 4.8% for F1, 7.2% for F2, 11.0% for F3, and 21.3% for F4. Considerable overlap of CPA values between Metavir stages was also observed (figure 5.1).

During follow up, 26 (4.9%) patients developed HCC, 39 (7.3%) developed liver decompensation and 33 (6.2%) had a liver related death. The AUROC for Metavir stage was 0.88 (95% CI, 0.82-0.95) for predicting HCC, 0.80 (95% CI, 0.71-0.88) for liver decompensation and 0.82 (95% CI, 0.74-0.90) for liver related death. Compared to Metavir stage CPA achieved a significantly higher AUROC of 0.92 (95% CI, 0.89-0.94) for predicting HCC (p=0.0112) and a similar AUROC of 0.79 (95% CI, 0.71-0.88) and 0.84 (95% CI, 0.76-0.93) for
decompensation and liver related death respectively. No significant improvement of AUROC was observed when CPA and Metavir were combined. 58 patients had the composite end-point of liver related death, HCC or liver decompensation during follow up. The AUROC for CPA to predict the composite end point was 0.83 (95% CI, 0.76-0.89). A cut point of 5% had a sensitivity of 87.9% and specificity of 51.4% in predicting the composite end point, whereas a cut point of 10% had a sensitivity of 62.1% and specificity of 90.3%, and a cut point of 20% had a sensitivity of 34.5% and specificity of 98.7% respectively. Using these cut points, CPA values were categorised into four stages: C1: 0%-5%, C2: 5%-10%, C3: 10%-20%, C4: >20%. 251 patients had C1, 201 had C2, 56 had C3 and 25 had C4.

Univariate analysis found that CPA stage, Metavir stage and age were significantly associated with liver decompensation, HCC development and liver related death (table 5.2). Multivariate analysis found that CPA stage and Metavir stage were independently associated with liver related death; CPA stage, Metavir stage and age were significantly associated with HCC development and Metavir stage was significantly associated with liver decompensation (table 5.2).

There was a significant difference in composite end point free survival between C1 and C2 (p=0.01), C2 and C3 (p<0.001), C3 and C4 (p<0.001) (figure 5.2). The 15 year composite end point free survival probability was 97.1% (95% CI, 94.0%-98.6%) for C1, 88.7% (95% CI, 80.6%-93.5%) for C2, 60.5% (95% CI, 45.2%-75.7%) for C3, 7.3% (95% CI, 0.5%-27.3%) for C4 (table 5.3). A significant difference was also found in separate analyses for almost all clinical
outcomes between each consecutive CPA stage (figure 5.2, table 5.3). Significance differences for HCC free survival were found between C1 and C2 (p=0.02), C2 and C3 (p<0.001), C3 and C4 (p=0.004) and for liver decompensation free survival between C2 and C3 (p=0.01), C3 and C4 (p<0.001) and for liver related survival between C2 and C3 (p<0.001) and C3 and C4 (p<0.001) (figure 5.2, table 5.3). Analysis of consecutive Metavir stages found a significant difference between F4 and F3 and F3 and F2 for the composite end point as well as all three end points. No significant difference in survival was found between F0 and F1 and F1 and F2 (figure 5.3).

The age and gender adjusted hazard ratio (HR) for decompensation free survival of C3 patients compared to those with C1-C2 was 3.6 (95% CI, 1.4-8.8), for HCC-free survival was 12.1 (95% CI, 3.9-37.4), for liver related survival was 7.2 (95% CI, 2.6-19.6). The age and gender adjusted HR for decompensation free survival of C4 patients compared to those with C1-C2 was 20.4 (95% CI, 9.0-46.4), for HCC-free survival was 36.7 (95% CI, 10.9-124), for liver related survival was 49.5 (95% CI, 19.6-125). Among the 39 cirrhotic patients, 4 (10.3%) had stage C2, 11 (28.2%) had C3 and 24 (61.5%) had stage C4. 7.1% of C2 and C3 patients had a liver related death during follow up compared with 60% of those with C4 (P<0.001). 20.0% of C2 and C3 patients and 41.7% of C4 patients developed HCC. 33.3% of C2 and C3 patients and 58.3% of C4 patients developed decompensation, but neither achieved a significant difference.

To test the influence of the size of the biopsy measurement area on the predictive ability of CPA stage, the interaction between CPA stage and biopsy
measurement area were included in Cox regression analysis. The interaction between CPA stage and measurement area was not significantly associated with any of the three end points suggesting that there was no significant influence of the size of the measurement area on the predictive ability of CPA values. To test the potential bias caused by exclusion of patients with successful treatment, a sensitivity analysis was conducted using the whole cohort (n=636) and similar results were found (data not shown).
Figure 5.1. Overlap of CPA values and Metavir stage. •: outliers
Figure 5.2. Clinical outcomes according to CPA stage. (A): composite end point free survival, (B): liver related survival, (C): decompensation free survival and (D): HCC free survival for each CPA stage.
Figure 5.3. Clinical outcomes according to Metavir stage. (A): composite end point free survival, (B): liver related survival, (C): decompensation free survival and (D): HCC free survival for each Metavir stage.
Table 5.1. Patients’ characteristics (n=533)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (SD)</td>
<td>41 (10)</td>
</tr>
<tr>
<td>Mean follow up years (range)</td>
<td>10.5 (0.1-20.2)</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>377 (70.7%)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>156 (29.3%)</td>
</tr>
<tr>
<td>Median CPA (range)</td>
<td>5.2 (1.3-44.6)</td>
</tr>
<tr>
<td>F0/F1/F2/F3/F4 (%)</td>
<td>22/47/16/8/7</td>
</tr>
<tr>
<td>C1/C2/C3/C4 (%)</td>
<td>47/38/11/5</td>
</tr>
<tr>
<td>HCC development, n (%)</td>
<td>26 (4.9)</td>
</tr>
<tr>
<td>Decompensation, n (%)</td>
<td>39 (7.3)</td>
</tr>
<tr>
<td>Liver related death, n (%)</td>
<td>33 (6.2)</td>
</tr>
</tbody>
</table>
Table 5.2. Predictors for decompensation, HCC and liver related death

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Decompensation</th>
<th>HCC</th>
<th>Liver related death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate</td>
<td>Multivariate</td>
<td>Univariate</td>
</tr>
<tr>
<td>age</td>
<td>1.07 (1.04-1.10)</td>
<td>ns</td>
<td>1.12 (1.08-1.15)</td>
</tr>
<tr>
<td>Metavir stage</td>
<td>3.04 (2.34-3.96)</td>
<td>2.09 (1.35-3.24)</td>
<td>4.50 (3.06-6.64)</td>
</tr>
<tr>
<td>CPA stage</td>
<td>3.78 (2.74-5.21)</td>
<td>ns</td>
<td>6.57 (4.17-10.35)</td>
</tr>
</tbody>
</table>

**Note:** ns: not significant

Table 5.3. 15 year survival probability of each end point for each CPA stage

<table>
<thead>
<tr>
<th>CPA stage</th>
<th>15 year survival probability (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Composite end point free survival</td>
</tr>
<tr>
<td>C1</td>
<td>97.1% (94.0%-98.6%)</td>
</tr>
<tr>
<td>C2</td>
<td>88.7% (80.6%-93.5%)</td>
</tr>
<tr>
<td>C3</td>
<td>60.5% (45.2%-75.7%)</td>
</tr>
<tr>
<td>C4</td>
<td>7.3% (0.5%-27.3%)</td>
</tr>
</tbody>
</table>
This study performed CPA measurements on a large and well documented CHC population that included all Metavir fibrosis stages. Normal human liver is estimated to contain approximately 5.5 mg collagen/g liver, and cirrhotic liver contains approximately 30 mg collagen/g liver (68). Animal studies found similar proportional increases in total extracellular matrix components from 5% in the normal liver to 25–40% in micronodular cirrhosis in rats (251). There was a significant correlation between CPA and Metavir stage \( (r = 0.738) \) and this was consistent with previous studies (75, 76, 238, 239). CPA values ranged from 1.3% to 44.6%. The median CPA in F0 (no fibrosis) patients was 3.7% which represents CPA values found in normal liver. CPA values found in cirrhosis (F4) ranged from 5.9% to 44.6% and the large range suggested that CPA had the potential to provide additional prognostic information in this group of patients. The continuous nature of the CPA measurement allowed the development of the optimal range of values that stratified different risks of adverse clinical outcomes. CPA values were consequently grouped into the following stages: C1: 0%-5% (normal), C2: 5%-10% (minimal), C3: 10%-20% (moderate), C4: >20% (severe). According to the new CPA staging system, 16% of patients had stage C3 or C4 which indicated significant liver fibrosis. This proportion was similar to those 15% of patients with Metavir stage F3 or F4. 64% of cirrhotic patients were staged C4 and potentially had the highest risk of developing adverse clinical outcomes. The considerable overlap between the two staging systems was not surprising as CPA stage and Metavir stage rely on different aspects of liver fibrosis progression. Metavir stage mainly depends on the severity of architectural distortion of the liver whereas CPA stage quantitates the amount of liver collagen.
This is the largest study to date that has validated the predictive ability of CPA to predict adverse clinical outcomes in CHC patients with all stages of fibrosis. Other studies have evaluated the utility of CPA in terms of risk stratification in those with cirrhosis or following liver transplantation (81-84). CPA had excellent predictive ability for HCC development with an AUROC of 0.92 which was significantly higher than that of Metavir stage. Moreover, CPA stage was able to stratify risk better than Metavir stage with a significant difference in HCC free survival between each consecutive CPA stage. CPA values and Metavir stage had a similar ability to predict liver related death and liver decompensation comparing AUROC values. There was a significant increase in liver related death and liver decompensation comparing CPA stage C2 to C3 and C3 to C4. Similar findings were found for Metavir stage F2 to F3 and F3 to F4. CPA and Metavir stage measure different features of liver fibrosis so it was not surprising that CPA stage remained an independent predictor for liver related death and HCC after adjusting for Metavir stage and age. CPA stage was also significantly associated with the risk of liver related death in cirrhotic patients. There was a trend for increased risk with CPA stage for HCC and liver decompensation but this was not significant and was possibly due to the limited number of cirrhotic patients. These findings support the hypothesis that CPA stage provides additional predictive value in addition to Metavir stage.

Due to the study design, CPA was measured using newly cut tissue slides from stored tissue blocks and this lead to a significantly smaller biopsy size for CPA measurement compared to Metavir staging (mean 10mm vs 15mm). Five mm was used as the minimum acceptable size for CPA analysis and inclusion of these small biopsies still resulted in an excellent ability of CPA to predict clinical
outcomes. No other study has tested the predictive ability of CPA using small liver biopsies. One study conducted by Bedossa et al suggested that an adequate biopsy length was 40mm for CPA measurement (78). The improved CPA measurement from smaller biopsy sizes was partly due to the exclusion of large portal tracts and capsule and also to the use of a high resolution scanner. Compared to histopathological staging, CPA measurement had additional requirements that include: a high resolution scanner, image analysis software and around five minutes analysis time for each biopsy. However, the CPA measurement is semi-automated and easily performed and requires less expertise and has a higher reproducibility than Metavir staging.

In summary, this study used a simple semi-automated slide scanning technique that measured collagen quantity in liver biopsies from CHC patients and this measurement was correlated with long term clinical outcomes. Moreover, CPA measurement has the potential to use biopsies 5mm or more in size. More studies are needed to validate the utility of CPA in other causes of liver disease.
6.1. ABSTRACT

6.1.1. Background and aims

Collagen proportional area (CPA) is a validated quantitative measurement of liver biopsy collagen and is measured using digital image analysis. This study aimed to develop a serum model that accurately predict CPA value.

6.1.2. Methods

213 chronic hepatitis C (CHC) patients were included and randomised into a training and validation set with 2:1 ratio. A CPA value was obtained for each biopsy using image analysis. Eleven serum markers were included in the modelling process.

6.1.3. Results

142 patients were included in the training set and 71 in the validation set. In the training set, univariate analysis found that hyaluronic acid (HA), gamma-glutamyl transpeptidase (GGT), α2-macroglobulin, platelet count, INR, prothrombin time, AST and age were significantly correlated with CPA value.
HA had the best correlation with a correlation coefficient value of 0.62. Three serum markers (HA, α2-macroglobulin and platelet count) which remained significant in multivariate analysis were included in the final model (CPAscore) and achieved an R square value of 0.46 to predict CPA. The mean CPAscore value was 7.70 (range: 0.98-28.2) and the mean variance between the CPAscore and measured CPA was 2.78. CPAscore had an AUROC of 0.86 (95% CI, 0.78-0.95) to predict those patients with a CPA ≥ 10% and a cut point of 8.7 had a sensitivity of 80.8% and specificity of 85.2%. The AUROC of CPAscore to predict patients with a CPA ≥ 20% was 0.96 (95% CI, 0.91-1.00) and a cut point of 10.7 had a sensitivity of 100% and specificity of 89%. A similar predictive ability of the final model was found in the validation set.

6.1.4. Conclusions

This study has for the first time developed a serum biochemical model using CPA as the reference standard. The model has the potential to improve the prediction of liver related clinical outcomes and non-invasively measures liver collagen.
6.2. INTRODUCTION

Liver fibrosis severity closely associates with progression of chronic liver disease and provides important information for patient management and prognosis. Liver biopsy is the reference standard to access liver fibrosis but is limited by its invasive manner, sampling error and risk of serious complications or death (237). As a result it has no longer been used as a clinical prognostic tool. In the last two decades, a large number of serum models were developed to predict liver fibrosis severity. Most of the models have been validated and showed satisfactory ability to predict significant fibrosis (Metavir F2-F4) with area under ROC curves (AUROC) varied from 0.77 to 0.87 (252). However, all currently in used serum models were built using semi-quantitative histopathological staging system as the reference standard. The limitations of histopathological stages such as inter and intra-observer variability of interpretation and small number of fibrosis severity categories would unavoidably decrease the accuracy of those serum models to predict fibrosis severity and the sensitivity to detect fibrosis progression or regression.

Quantitative image analysis of liver biopsy specimens is based on the analysis of digital images that measures the area of collagen and total area of liver tissue and then calculates the proportion of the area of the biopsy occupied by collagen [collagen proportional area (CPA)] (248). Unlike semi-quantitative staging systems which mainly depend on the degree of architectural distortion, image analysis can measure the amount of fibrosis in a linear and continuous scale. Several studies have shown a significant correlation between CPA and histological stage (75, 76, 133, 239). CPA also correlated with hepatic venous pressure gradient in CHC patients (76), with decompensation in cirrhotic
patients (81), with transplantation free survival in children with biliary atresia (249) and post liver transplantation outcomes for recurrent hepatitis C virus (HCV) (82-84). Our previous studies demonstrated a high accuracy of CPA to predict long term clinical outcomes. A CPA value of 10% to 20% and a CPA value ≥20% were associated with moderate and high risk level of adverse clinical outcomes respectively (253). This suggested that CPA has the potential to better illustrate the severity of liver fibrosis than histopathological staging system.

The fact that CPA and serum markers are both continuous variables means that CPA potentially has better correlation with serum markers than histopathological stages. However, few studies associated serum markers with CPA values. An early study found CPA was better correlated with serum markers than histopathological stages of liver fibrosis (75). Hyaluronate and prothrombin time had the best correlation with CPA (75). Another study found TIMP-1, ALP ratio and platelet count significantly correlated with CPA, but they transformed CPA to a categorical variable (upper 31%, lower 69% of CPA values) which might diminish its advantage as a continuous measurement. Our previous study also identified eight serum markers that was significantly correlated with CPA values (248). Up to today, no effort was done to develop a serum model to predict CPA value. This study aimed to build a serum model to detect liver fibrosis severity using CPA as the reference standard.
6.3. METHODS

6.3.1. Patients

Chronic hepatitis C patients from the liver clinics at Sir Charles Gairdner Hospital who had a liver biopsy and a Hepascore (122) performed within six months of biopsy from January 1997 until January 2012 were included. Exclusion criteria included co-infection with HBV and HIV; other liver diseases including hemochromatosis, α1-antitrypsin deficiency, Wilson’s disease and autoimmune liver diseases; successful anti-viral treatment before biopsy; previous liver transplantation. The study was approved by the Sir Charles Gairdner Hospital Human Research Ethics Committee.

6.3.2. Image analysis of liver fibrosis

A new liver section was cut from the stored tissue block and stained using sirius red and this was used for image analysis. CPA was measured using the optimised method previously described (248). Sections were scanned using the Aperio Scanscope XT Digital Slide Scanner at 20X magnification (1.59X10^7 Pixels = 1 mm^2). The image was viewed using Aperio ImageScope software version 10.0. The liver capsule and large portal tracts were excluded as these did not represent disease related collagen (172). The optimum threshold for positive pixels that corresponded to areas of sirius red staining was determined according to hue value and colour saturation using the original image for comparison. A binary image was produced by the software and CPA was expressed as a percentage of positive pixels to total pixels. The CPA value was calculated by the software. According to our previous study, a biopsy with a measurement area less than 5 mm^2 after exclusion of any large portal tracts or
capsule was considered as insufficient and thus excluded from further analysis (248).

6.3.3. Clinical information

Eleven serum markers namely: hyaluronic acid (HA), bilirubin, gamma-glutamyl transpeptidase (GGT), α2-macroglobulin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelet count, prothrombin time, international normalized ratio (INR), alkaline phosphatase (ALP) and albumin were obtained from hospital database. For patients had multiple blood tests, result that closest to liver biopsy date were recorded. All serum maker test dates were within six months of liver biopsy date. Patients’ age, gender and Metavir fibrosis stage were also recorded.

6.3.4. Statistical analysis

Included patients were randomised into two groups: a training set and a validation set in a 2:1 ratio. Only patients in the training set were included in the modelling process. Continues variables were expressed as mean, standard deviation (SD) and mean values were compared using t test. Categorical variables were expressed as count, percentage and proportions were compared using Chi-square test. Correlations between CPA and serum markers were determined by Pearson correlation coefficients. Variables that have a significant correlation with CPA were included in the modelling process. Linear regression analysis was used to determine the optimal model to predict CPA. AUROC was calculated for the new models to predict CPA larger than 10% and 20%. Cut
points were chosen based on Youden’s index. The predictive ability of the final model was also evaluated in the validation set.
6.4. RESULTS

213 patients were included: 142 patients were randomised into training set and 71 were randomised into validation set (table 6.1). Both groups showed widespread of liver fibrosis severity with wide distribution of Metavir stage as well as CPA values. CPA ranged from 1.6% to 32.67% in training set and from 2.77% to 21.3% in validation set. No significant difference was found concerning Metavir stage, CPA value and serum markers between the two groups.

6.4.1. Training set

Eleven serum markers were included in univariate analysis, eight of them (HA, GGT, α2-macroglobulin, platelet count, INR, prothrombin time, albumin and AST) were significantly correlated with CPA value (table 6.2). HA had the best correlation with CPA among all included serum markers with a correlation coefficient of 0.62. Age was significant correlated with CPA with correlation coefficient of 0.329, p<0.0001. No significant association was found between gender and CPA. Variables that showed significant correlations with CPA were further included into multivariate analysis. Model including all nine variables achieved an R square value of 0.52. The model that included four variables (HA, α2-macroglobulin, platelet count and albumin) had an R square value of 0.47 and the model that included three variables (HA, α2-macroglobulin, platelet count) achieved an R square value of 0.46 (table 6.3). The final model (CPAscore) was chosen as 0.026* HA (ug/L) + 0.620 * macroglobulin (g/L) - 0.019 * platelet count (X10^9/L) + 8.06. HA, α2-macroglobulin and platelet count showed independent correlation with CPA value with a p value of <0.001, 0.069
and 0.001 respectively. CPAscore was calculated for each patient, with a mean value of 7.70 (range: 0.98-28.2). As showed in figure 6.1, most actual CPA value located around the line that representative of CPAscore with the mean difference between actual CPA value and CPAscore of 2.78. Two clinical relevant cut points of CPA value were defined in previous study (253). A CPA value between 10% and 20% and a CPA value larger than 20% were associated with moderate and high risk of adverse clinical outcomes respectively (253). CPAscore achieved an AUROC of 0.86 (95% CI, 0.78-0.95) to predict patient with CPA ≥10% and cut point of 8.7 had a sensitivity of 81% and specificity of 85%. The AUROC of the model to predict patients with CPA ≥20% was 0.96 (95% CI, 0.91-1.00) and cut point of 10.7 achieved a sensitivity of 100% and specificity of 89% (table 6.3, figure 6.2).

6.4.2. Validation set

The predictive ability of CPAscore was then tested in validation set. This group had a mean actual CPA value of 7.83% and a mean CPAscore value of 7.12. Similar correlation between CPAscore and actual CPA value was observed with a correlation coefficient of 0.63 (p<0.001) (figure 6.1). This model achieved an AUROC of 0.82 (95% CI, 0.70-0.93) to predict CPA ≥10% and an AUROC of 0.94 (95% CI, 0.89-1.00) to predict CPA ≥ 20%. The cut point of 8.7 had a sensitivity of 65% and specificity of 89% to detect patients with CPA ≥10% and the cut point of 10.7 achieved a sensitivity of 100% and specificity of 93% to detect patients with CPA ≥ 20%.
Figure 6.1. Correlations between CPAscore and actual CPA value.
Figure 6.2. ROC curves for CPAscore. (A) to predict CPA ≥10% and (B) to predict CPA ≥20%.
### Table 6.1. Patient characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Training set (n=142)</th>
<th>Validation set (n=71)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>42 (10)</td>
<td>44 (11)</td>
<td>0.3678</td>
</tr>
<tr>
<td>Gender M/F</td>
<td>106/36</td>
<td>57/14</td>
<td>0.396</td>
</tr>
<tr>
<td>HA (ug/L)</td>
<td>65 (93)</td>
<td>49 (52)</td>
<td>0.1912</td>
</tr>
<tr>
<td>GGT (U/L)</td>
<td>81 (74)</td>
<td>100 (119)</td>
<td>0.1509</td>
</tr>
<tr>
<td>Bilirubin (umol/L)</td>
<td>11 (5)</td>
<td>11 (5)</td>
<td>0.4729</td>
</tr>
<tr>
<td>α2-macroglobulin (g/L)</td>
<td>2.98 (1.19)</td>
<td>2.85 (1.16)</td>
<td>0.4429</td>
</tr>
<tr>
<td>Platelet count (10^9/L)</td>
<td>212 (70)</td>
<td>216 (66)</td>
<td>0.7053</td>
</tr>
<tr>
<td>INR</td>
<td>1 (0.1)</td>
<td>1.0 (0.1)</td>
<td>0.9581</td>
</tr>
<tr>
<td>PT (secs)</td>
<td>8.7 (1.6)</td>
<td>8.6 (1.0)</td>
<td>0.5527</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>122 (91)</td>
<td>122 (90)</td>
<td>0.9828</td>
</tr>
<tr>
<td>AST (U/L)</td>
<td>78 (51)</td>
<td>76 (63)</td>
<td>0.7701</td>
</tr>
<tr>
<td>ALP (U/L)</td>
<td>85 (62)</td>
<td>85 (40)</td>
<td>0.9487</td>
</tr>
<tr>
<td>Albumin (g/L)</td>
<td>44 (15)</td>
<td>43 (3)</td>
<td>0.3258</td>
</tr>
<tr>
<td>CPA (%)</td>
<td>7.66 (5.42)</td>
<td>7.83 (4.54)</td>
<td>0.8238</td>
</tr>
<tr>
<td>F0/F1/F2/F3/F4</td>
<td>22/55/34/20/11</td>
<td>12/27/13/11/8</td>
<td>0.823</td>
</tr>
</tbody>
</table>

**Note:** Continuous variables were presented as mean (standard deviation). Categorical variables were presented as count.

### Table 6.2. Correlations between serum markers and CPA.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Correlation</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA</td>
<td>0.624</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>GGT</td>
<td>0.198</td>
<td>P=0.018</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>0.023</td>
<td>p=0.783</td>
</tr>
<tr>
<td>α2-macroglobulin</td>
<td>0.458</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Platelet count</td>
<td>-0.477</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>INR</td>
<td>0.418</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>PT</td>
<td>0.328</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>ALT</td>
<td>0.043</td>
<td>p=0.618</td>
</tr>
<tr>
<td>AST</td>
<td>0.267</td>
<td>p=0.004</td>
</tr>
<tr>
<td>ALP</td>
<td>0.101</td>
<td>p=0.247</td>
</tr>
<tr>
<td>albumin</td>
<td>-0.352</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Coefficients of variables, p value</td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>HA</td>
<td>0.019, p=0.001</td>
<td>0.018, p=0.001</td>
</tr>
<tr>
<td>α2-macroglobulin</td>
<td>0.734, p=0.048</td>
<td>0.649, p=0.067</td>
</tr>
<tr>
<td>platelet count</td>
<td>-0.019, p=0.001</td>
<td>-0.019, p=0.001</td>
</tr>
<tr>
<td>albumin</td>
<td>-0.279, p=0.017</td>
<td>-0.267, p=0.021</td>
</tr>
<tr>
<td>INR</td>
<td>7.00, p=0.070</td>
<td>6.78, p=0.086</td>
</tr>
<tr>
<td>age</td>
<td>-0.034, p=0.429</td>
<td>-</td>
</tr>
<tr>
<td>constant</td>
<td>15.33</td>
<td>13.43</td>
</tr>
<tr>
<td>R square</td>
<td>0.48</td>
<td>0.48</td>
</tr>
<tr>
<td>AUROC for CPA≥10% (95% CI)</td>
<td>0.86 (0.78-0.95)</td>
<td>0.87 (0.79-0.95)</td>
</tr>
<tr>
<td>AUROC for CPA≥20% (95% CI)</td>
<td>0.98 (0.95-1.00)</td>
<td>0.98 (0.95-1.00)</td>
</tr>
</tbody>
</table>
6.5. DISCUSSION

This is the first attempt to build a serum model to predict CPA values. Compared to histopathological staging system, CPA quantitatively measures liver fibrosis severity. This allows CPA to have a more direct correlation with serum markers and simplifies the interpretation of the values of accordingly built models. Model developed using CPA as a reference standard has the potential to have higher accuracy to detect fibrosis changes. This is of great clinical importance in the new era of antiviral therapy where fibrosis regression becomes one of the major goals of treatment.

CPAscore closely associated with CPA value with a correlation coefficient of 0.63. The model fit (R square) of the CPAscore was 0.46, suggesting CPAscore was not the only factor that reflected CPA value. This could be partly explained by the continuous nature of CPA value and serum marker values, both of which may include certain background variations that didn't represent the actual fibrosis severity. Individual variation was one important potential variation and it was possible that CPAscore would have higher accuracy to detect fibrosis change for single individual by using multiple sequential tests. Currently, there is no non-invasive method that has the acceptable accuracy to detect fibrosis progression and regression during antiviral treatment. CPAscore proposed in this study has the potential usage in clinical practice to select most suitable patients for antiviral treatment and to evaluate the efficacy of the treatment. Moreover, this model was advantageous for its simplicity including only HA, α2-macroglobulin and platelet count. The tests of these three serum markers were easy and readily available in many medical centres.
A number of serum models were developed to detect fibrosis severity using histopathological staging systems as the reference standard. A recent study reviewed 20 serum models and reported the AUROC of those models to predict significant fibrosis (Metavir F2-F4) ranged from 0.77 to 0.87 with the highest accuracy achieved by FibroTest and the AUROC to predict cirrhosis (Metavir F4) ranged from 0.81 to 0.89 with the highest accuracy achieved by APRI, ELF and Hepascore (252). The different modelling process made it hard to directly compare the predictive ability of CPAscore with that of other existing serum models. Two cut points of CPA value were determined in our previous study. Patient with CPA value between 10% and 20% had moderate risk of developing liver related death, hepatocellular carcinoma and liver decompensation with 15 year survival probability of 68.7%, 68.3% and 74.2% respectively (253). Patients with CPA larger than 20% had a significant higher risk with corresponding 15 year survival probability of 31.8%, 24.8% and 25.7% respectively (253). The CPAscore achieved satisfying accuracy to detect those patients with CPA larger than 10% and those larger than 20% with AUROC of 0.86 and 0.96 respectively in training set and 0.82 and 0.94 respectively in validation set. Using cut points of 8.7 and 10.7, patients with CPA larger than 10% and 20% were detected respectively with both high sensitivity and specificity. This suggested that CPAscore had higher accuracy to stratify patients into low, moderate and high risk group than other existing models.

In conclusion, this study for the first time developed a serum model using CPA as the reference standard. This model had both the ability to quantitatively represent liver fibrosis severity and high accuracy to detect patients with moderate and severe liver fibrosis for CHC patients. More studies are needed to
validate this model in patients with different types of chronic liver disease and to compare it with existing models.
CHAPTER 7

NEW SIMPLE SERUM MARKER MODELS RELIABLY PREDICT LIVER RELATED CLINICAL OUTCOMES IN CHRONIC HEPATITIS C INFECTION

7.1. ABSTRACT

7.1.1. Background and aims

Few serum models have been developed to directly predict clinical outcomes among patients with chronic hepatitis C (CHC). This study aimed to develop serum models that directly predict liver related death, hepatocellular carcinoma (HCC) and liver decompensation respectively.

7.1.2. Methods

617 CHC patients with serum markers available were randomized into a training set (n=411) and a validation set (n=206). Clinical outcomes were determined using population based data-linkage system.

7.1.3. Results

Mean follow up was six years during which 22 liver related death, 23 HCC and 27 liver decompensation were observed. Using the training set, albumin, GGT, HA, age and sex were chosen in the final model to predict five year liver related death with AUROC of 0.95. Using two cut points of 4.0 and 5.5, patients were
categorised into three risk groups with an annual incidence rate for liver related death of 0.1%, 2% and 13.2% respectively (p<0.001). Albumin, GGT, HA, age and sex were used to predict five year liver decompensation with AUROC of 0.90. A cut point of 4.5 achieved a sensitivity of 94% and a specificity of 84% to predict five year decompensation. Using this cut point patients were divided into two risk groups with an annual incidence rate for decompensation of 0.2% and 5.8% respectively (p<0.001). ALP, α2-macroglobulin, age and sex were chosen to predict five year HCC occurrence with AUROC of 0.95. A cut point of 8 had a sensitivity of 90% and specificity of 88% to predict 5 year HCC occurrence. Using this cut point patients were divided into two risk groups with an annual incidence rate for HCC of 0.2% and 5.6% respectively (p<0.001). Similar results were obtained using the validation set.

7.1.4. Conclusions

All three simple models had excellent predictive accuracy and were able to stratify risk into clinical meaningful categories for CHC patients.
7.2. INTRODUCTION

Hepatitis C virus (HCV) infection affects about 180 million people worldwide and predisposes these patients to complications of cirrhosis, hepatocellular carcinoma and early death (225). Epidemiological analyse showed that patients with chronic hepatitis C (CHC) had a three times higher risk of overall death and a 17 times higher risk of liver related death than the general population (226, 227). However, identifying those CHC patients who are at higher risk of developing liver related morbidity and mortality is problematic. This is due to the variable natural history of HCV with its prolonged and predominantly asymptomatic early phase and variable later progression. The histopathological stage of fibrosis has long been used to further stratify risk in CHC patients (214, 215). However, liver biopsy is not routinely performed for prognostic purposes as it is an invasive procedure with sampling error and risk of serious complications (237).

Two types of non-invasive clinical tests are available to predict the severity of liver fibrosis and these are serum panel tests and radiological tests such as transient elastography and MR elastography (247). However, only a few of the serum panels have been shown to be associated with liver related clinical outcomes. A study in patients with alcoholic liver disease found that Fibrotest, Hepascore and Fibrometer all had a moderate ability to predict liver related death with a median follow up of eight years (254). Another study compared the prognostic ability of serum markers (FibroMeter, CirrhoMeter, Fibrotest, Hepascore, FIB4 and APRI) with that of Metavir stages and found higher accuracy of those serum models to predict liver related death as well as liver related events in CHC patients (255). A Meta-analysis of six studies that
included patients with all types of chronic liver disease found that Fibrotest predicted five year liver related death better than FIB-4 or APRI (216). Hepascore achieved an AUROC of 0.86 to predict liver related death (217). The ELF test also has been associated with five year liver related outcomes in chronic liver disease (211). Studies showed that transient elastography is significantly correlated with the development of liver decompensation, hepatocellular carcinoma (HCC), death and composite of these outcomes for patients with chronic liver disease (213, 223). This evidence suggested the potential ability of non-invasive methods for fibrosis assessment to predict clinical outcomes.

The potential advantages of serum panels that have been developed to directly predict clinical outcomes are that they will incorporate additional analytes that are not useful in predicting fibrosis but will be useful in predicting clinical end points. These additional analytes may be associated with other factors such as portal hypertension, coagulopathy, protein synthetic dysfunction and renal failure that are known to predict liver related outcomes. The MELD score was the first serum model developed to predict short term clinical outcomes following the TIPPS procedure (190). A modified version is now widely used for prioritisation of organ allocation in patients on the liver transplantation waitlist. It has a high accuracy to predict short term mortality (three month) for these patients, but its predictive accuracy decreases significantly if the target population is extended to all cirrhotic patients, or for predicting mid or long term outcomes (191, 197).
The aim of this study was to develop simple serum panel models that directly predict the risk of liver related death, HCC and decompensation in a large group of CHC patients who had long term clinical follow up data available. The new direct models were compared with those models that were previously developed to predict liver fibrosis.
7.3. METHODS

7.3.1. Patients

CHC patients who attended the Department of Gastroenterology and Hepatology, Sir Charles Gairdner Hospital and had a Hepascore and other serum markers available from 1997 to 2012 were included. CHC was defined at a positive HCV RNA detected by PCR on two occasions greater than six months apart. Exclusion criteria included co-infection with hepatitis B virus and human immunodeficiency virus; other liver diseases including hemochromatosis, α1-antitrypsin deficiency, Wilson disease and autoimmune liver diseases; previous liver transplantation; and episodes of liver decompensation and HCC before the inclusion date. The study was approved by the Sir Charles Gairdner Hospital Human Research Ethics Committee and the Western Australia Department of Health Human Research Ethics Committee.

7.3.2. Clinical data

Hepascore and its four component serum markers: hyaluronic acid (HA), bilirubin, gamma-glutamyl transpeptidase (GGT), α2-macroglobulin and other serum markers: alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelet count, prothrombin time, international normalized ratio (INR), alkaline phosphatase (ALP), creatinine and albumin were included in the study. Patients’ age and sex were also recorded. The APRI was calculated as AST (U/L)/upper normal * 100/platelet count (10^9/L) (120). The Lok index was calculated as -5.56-0.0089 * platelet (10^3/mm^3) +1.26 * AST/ALT+5.27 * INR (129). The FIB-4 was calculated as age (years) * AST [U/L]/ ((platelets [10^9/L] * (ALT [U/L])1/2) (123). The MELD score was calculated as 10+ (0.957*
\[
\ln(\text{creatinine (umol/L)/88.4}) + 0.378 \times \ln(\text{bilirubin (umol/L)/ 17.1}) + 1.12 \times \ln(\text{INR})) +6.43 \ (190).
\]

7.3.3. Clinical outcomes

The long term follow up of patients was obtained from the Western Australian Data Linkage Unit. This is a validated population-based data linkage system that links multiple health related datasets including the state cancer register, the state hospital morbidity database and the state mortality records dating back to 1982, 1970 and 1969 respectively (229). The Hospital Morbidity Data System has 100% coverage of data for hospital admissions throughout the state with a record linkage success rate larger than 99% (229). The hospital admission diagnosis and the cause of death were recorded using ICD 9 (before 1997) and ICD 10 (after 1997) classification codes. Personal identifiers were encrypted and stored separately from the data used for analysis.

7.3.4. Endpoints and statistical analysis

The primary endpoint was liver related death or liver transplantation. Liver related death was defined as death from liver failure, variceal bleeding or hepatocellular carcinoma as well as death in which liver disease was the major contributing factor. The secondary endpoints were the first episode of liver decompensation, (ascites, hepatic encephalopathy, variceal bleeding, hepatorenal syndrome, spontaneous bacterial peritonitis) or the development of HCC. Patients were followed from the date of serum tests until the occurrence of an end point or the end of the study.
Patients were randomised into two groups: a training set and a validation set in a 2:1 ratio. Cox regression analysis was used to model survival and predict liver related death, liver decompensation and HCC respectively. The candidate variables for each model were those factors that had a significant association with each end point with p value less than 0.05 using a univariate cox model. Age and sex were also included in all models as they are known factors that could affect both survival and serum marker results. The final models were chosen using the backwards selection method. ROC curve analysis was used to test the ability of the final model to predict three year, five year and ten year risk of each end point and cut points were defined using Youden index. The survival probability for each risk group within the new models was calculated using Kaplan-Meier curves and a significance difference was defined with the log rank test. AUROC was calculated for each of the new models and this was compared with APRI, FIB-4, Lok index, Hepascore and MELD score. The incident rate of end points were calculated and compared using the Z test.
7.4. RESULTS

617 patients were included: 411 patients were randomised to the training set and 206 to the validation set. Patients’ characteristics of the two groups are shown in table 7.1. Patient follow up was for a mean of six years (range 0.1-14.1). There were 22 liver related deaths or liver transplantations, 23 HCC’s and 27 episodes of decompensation by the end of follow up.

7.4.1. Training set

Liver related death was significantly associated with HA, GGT, bilirubin, α2-macroglobulin, platelet count, INR, prothrombin time, AST, ALP, albumin and creatinine. These were included with age and sex in the initial predictive model (table 7.2). The model with the highest AUROC to predict five year liver related death was chosen as the final model [Liver Outcome Score (LOS)]: LOS_death = \(-0.1792^*\text{albumin} + 0.0042^*\text{GGT} + 0.0041^*\text{HA} + 0.0377^*\text{age} + 0.4492\) (if sex=male) +8 (table 7.2). The AUROC for LOS_death to predict three year, five year and ten year liver related death was 0.96 (95% CI, 0.91-1.00), 0.95 (95%CI, 0.90-1.00) and 0.95 (95% CI, 0.91-0.99) respectively. The mean LOS_death value in the training set was 2.99 (range: 0.23-9.85). A cut point of 5.5 had a sensitivity of 80.0% and specificity of 96.5% to predict three year liver related death. A cut point of 4.0 had a sensitivity of 92.9% and specificity of 85.1% to predict ten year liver related death. Using these cut points, patients were categorised into low, moderate and high risk group (< 4.0, 4.0-5.5, 5.5) with the annual incident rate of 0.1%, 2.02% and 13.2% respectively (table 6.3). A significant difference of liver related survival was found between groups (P<0.001) (figure 7.1).
Liver decompensation was significantly associated with ten variables (HA, GGT, bilirubin, α2-macroglobulin, platelet count, INR, prothrombin time, AST, ALP, albumin) and these were included with age and sex as candidate variables in the initial model (table 7.2). The model with highest AUROC to predict five year decompensation was \[ \text{LOS}_{\text{decompensation}} = 0.0031 \times \text{HA} + 0.0030 \times \text{GGT} + 0.0562 \times \text{age} - 0.5342 \times \text{if sex=male} - 0.1870 \times \text{albumin} + 9 \] (table 7.2). The AUROC for LOS_{decompensation} to predict three year, five year and ten year decompensation was 0.96 (0.93-0.99), 0.90 (0.80-1.00) and 0.89 (95% CI, 0.80-0.98) respectively. The mean LOS_{decompensation} values in the training set was 3.59 (range: 0.39 - 9.23). A cut point of 4.5 achieved a sensitivity of 94.4% and a specificity of 83.5% to predict five year decompensation development. The same cut point had a sensitivity of 85.7% and specificity of 84.2% to predict ten year decompensation development and a sensitivity of 100% and specificity of 83.4% to predict three year decompensation development. Patients were therefore categorised into low and high risk group of developing decompensation with annual incidence rate of 0.15% and 5.58% respectively (table 7.3). Significant difference in decompensation free survival was found between these two groups (p<0.001) (figure 7.1).

The development of HCC was significantly associated with eleven variables (HA, GGT, bilirubin, α2-macroglobulin, platelet count, INR, prothrombin time, AST, ALT, ALP, albumin) and these were included with age and sex as candidate variables to predict HCC (table 7.2). The final model with the highest AUROC to predict five year HCC development was \[ \text{LOS}_{\text{HCC}} = 1.731 \times \text{if sex=male} + 0.0093 \times \text{ALP} + 0.6408 \times \alpha_2\text{-macroglobulin} + 0.1350 \times \text{age} - 4 \]. The AUROC for LOS_{HCC} to predict ten year, five year and three year HCC development was
0.94 (95% CI, 0.90-0.99), 0.95 (95% CI, 0.91-0.99 ) and 0.93 (95% CI, 0.89-0.98) respectively. The mean LOS_HCC value in the training set was 5.75 (range: 0.45-12.01). A cut point of 8 had a sensitivity of 90.0% and specificity of 87.9% to predict five year HCC. The same cut point had a sensitivity of 80% and specificity of 88.6% to predict ten year HCC development and a sensitivity of 88.9% and a specificity of 87.7% to predict three year HCC development. Patients were therefore categorised into low and high risk group of developing HCC with annual incidence rate of 0.15% and 5.78% respectively (table 7.3). Significant difference of HCC free survival was found between these two groups (p<0.001) (figure 7.1).

7.4.2. Validation set

Patients in the validation set were followed for a mean of six years. Seven patients had a liver related death, six had liver decompensation and seven developed HCC. In this group the AUROC of LOS_death to predict three year, five year and ten year liver related death was 0.94 (95% CI, 0.89-1.00), 0.96 (95% CI, 0.92-1.00) and 0.95 (95% CI, 0.91-0.99) respectively. The AUROC of LOS_decompensation to predict three year, five year and ten year decompensation was 0.94 (95% CI, 0.85-1.00), 0.95 (95% CI, 0.85-1.00) and 0.87 (95% CI, 0.76-0.99) respectively. The AUROC of LOS_HCC to predict three year, five year and ten year HCC was 0.92 (95% CI, 0.84-1.00), 0.93 (95% CI, 0.87-1.00) and 0.94 (95% CI, 0.90-0.99) respectively.
7.4.3. Comparison with other serum models

The predictive ability of the LOS panel was compared with MELD score, Hepascore, APRI, FIB-4 and the Lok index (table 7.4). The LOS panel had the best ability to predict liver related death, HCC and liver decompensation among serum models. Individual comparison showed that LOS_death was significantly better than Hepascore ($p=0.0009$) and MELD score ($p=0.0076$) to predict liver related death. LOS_decompensation was significantly better than MELD score ($p=0.0039$) to predict liver decompensation and LOS_HCC was significantly better than all other serum models to predict HCC development (figure 7.2).
Figure 7.1. Survival curves according to LOS panel. (A): Liver related survival (B): Decompensation free survival (C): HCC free survival.
Figure 7.2. ROC curves of serum models to predict HCC development.
Table 7.1. Patients characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Training set (n=411)</th>
<th>Validation set (n=206)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>44 (10)</td>
<td>43 (11)</td>
<td>0.366</td>
</tr>
<tr>
<td>Gender M/F</td>
<td>284/127</td>
<td>133/73</td>
<td>0.256</td>
</tr>
<tr>
<td>HA (ug/L)</td>
<td>77 (137)</td>
<td>74 (141)</td>
<td>0.745</td>
</tr>
<tr>
<td>GGT (U/L)</td>
<td>95 (122)</td>
<td>87 (106)</td>
<td>0.374</td>
</tr>
<tr>
<td>Bilirubin (umol/L)</td>
<td>12 (7)</td>
<td>11 (16)</td>
<td>0.746</td>
</tr>
<tr>
<td>α2-macroglobulin (g/L)</td>
<td>2.74 (1.08)</td>
<td>2.69 (1.03)</td>
<td>0.547</td>
</tr>
<tr>
<td>Platelet count (10^9/L)</td>
<td>220 (81)</td>
<td>224 (92)</td>
<td>0.557</td>
</tr>
<tr>
<td>INR</td>
<td>1.0 (0.1)</td>
<td>1.0 (0.1)</td>
<td>0.676</td>
</tr>
<tr>
<td>Creatinine (umol/L)</td>
<td>76.5 (1.7)</td>
<td>78.5 (3.01)</td>
<td>0.551</td>
</tr>
<tr>
<td>prothrombin time (secs)</td>
<td>8.9 (1.4)</td>
<td>8.7 (1.3)</td>
<td>0.066</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>117 (101)</td>
<td>119 (140)</td>
<td>0.844</td>
</tr>
<tr>
<td>AST (U/L)</td>
<td>76 (60)</td>
<td>78 (111)</td>
<td>0.723</td>
</tr>
<tr>
<td>ALP (U/L)</td>
<td>85 (51)</td>
<td>88 (45)</td>
<td>0.709</td>
</tr>
<tr>
<td>Albumin (g/L)</td>
<td>44 (9.5)</td>
<td>43 (3.9)</td>
<td>0.280</td>
</tr>
<tr>
<td>Follow up time (years)</td>
<td>6.0 (3.0)</td>
<td>6.2 (3.1)</td>
<td>0.429</td>
</tr>
<tr>
<td>Liver related death</td>
<td>15 (3.65%)</td>
<td>7 (3.40%)</td>
<td>0.874</td>
</tr>
<tr>
<td>HCC</td>
<td>16 (3.89%)</td>
<td>7 (3.40%)</td>
<td>0.760</td>
</tr>
<tr>
<td>Decompensation</td>
<td>21 (5.11%)</td>
<td>6 (2.91%)</td>
<td>0.255</td>
</tr>
</tbody>
</table>

Note: Continuous variables were presented as mean (standard deviation) and categorical variables were presented as count (percentage).
Table 7.2. Univariate and multivariate analysis to predict each end point

<table>
<thead>
<tr>
<th></th>
<th>p value</th>
<th>Liver related death</th>
<th>Decomposition</th>
<th>HCC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate</td>
<td>Multivariate</td>
<td>Univariate</td>
<td>Multivariate</td>
</tr>
<tr>
<td>HA</td>
<td>P&lt;0.001</td>
<td>P=0.004</td>
<td>P&lt;0.001</td>
<td>P=0.017</td>
</tr>
<tr>
<td>GGT</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P=0.005</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>P&lt;0.001</td>
<td>P=0.004</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>α2-macroglobulin</td>
<td>P=0.001</td>
<td>P=0.004</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Platelet count</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>INR</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>PT</td>
<td>P=0.002</td>
<td>P=0.010</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>ALT</td>
<td>P=0.073</td>
<td>P=0.915</td>
<td></td>
<td>P=0.011</td>
</tr>
<tr>
<td>AST</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Creatinine</td>
<td>P=0.008</td>
<td>P=0.681</td>
<td></td>
<td>P=0.857</td>
</tr>
<tr>
<td>ALP</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Albumin</td>
<td>P&lt;0.001</td>
<td>P=0.067</td>
<td>P&lt;0.001</td>
<td>P=0.034</td>
</tr>
<tr>
<td>Age</td>
<td>P=0.001</td>
<td>P=0.148</td>
<td>P&lt;0.001</td>
<td>P=0.019</td>
</tr>
<tr>
<td>Sex</td>
<td>P=0.353</td>
<td>P=0.516</td>
<td>P=0.484</td>
<td>P=0.267</td>
</tr>
</tbody>
</table>
Table 7.3. Incident rates for each endpoint according to corresponding models

<table>
<thead>
<tr>
<th>Clinical outcomes</th>
<th>Groups</th>
<th>Annual incident rate, 95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training set</td>
<td>Validation set</td>
</tr>
<tr>
<td>Liver related death</td>
<td>low risk</td>
<td>0.10% (0.04-0.24%)</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>moderate risk</td>
<td>2.02% (0.25-3.80%)</td>
<td>4.70% (0.62-10.1%)</td>
</tr>
<tr>
<td></td>
<td>high risk</td>
<td>13.2% (4.05-22.3%)</td>
<td>11.5% (0.23-22.8%)</td>
</tr>
<tr>
<td>Decompensation</td>
<td>low risk</td>
<td>0.15% (0.02-0.33%)</td>
<td>0.28% (0.04-0.59%)</td>
</tr>
<tr>
<td></td>
<td>high risk</td>
<td>5.58% (3.00-8.16%)</td>
<td>2.14% (0.28-4.57%)</td>
</tr>
<tr>
<td>HCC</td>
<td>low risk</td>
<td>0.15% (0.02-0.33%)</td>
<td>0.09% (0.00-0.28%)</td>
</tr>
<tr>
<td></td>
<td>high risk</td>
<td>5.78% (2.51-9.05%)</td>
<td>7.57 (1.51-13.62%)</td>
</tr>
</tbody>
</table>

Note: P value was calculated for comparing the annual incident rate between each risk group in the whole cohort.

Table 7.4. Comparison of the accuracy of LOS panel with other models in predicting liver outcomes

<table>
<thead>
<tr>
<th>Models</th>
<th>AUROC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Liver related death</td>
</tr>
<tr>
<td>LOS death</td>
<td>0.95 (0.92-0.97)</td>
</tr>
<tr>
<td>LOS decompenation</td>
<td>-</td>
</tr>
<tr>
<td>LOS HCC</td>
<td>-</td>
</tr>
<tr>
<td>Hepascore</td>
<td>0.87 (0.82-0.93)</td>
</tr>
<tr>
<td>APRI</td>
<td>0.89 (0.84-0.94)</td>
</tr>
<tr>
<td>FIB-4</td>
<td>0.92 (0.87-0.96)</td>
</tr>
<tr>
<td>Lok index</td>
<td>0.86 (0.75-0.96)</td>
</tr>
<tr>
<td>Meld score</td>
<td>0.76 (0.61-0.91)</td>
</tr>
</tbody>
</table>

Note: Models were tested in validation set.
7.5. DISCUSSION

The development of simple serum panel models that are able to stratify CHC patients into a hierarchy of risk levels of adverse clinical outcomes is of considerable clinical significance. The strengths of this study were the inclusion of a large number of well characterized CHC patients with a broad spectrum of disease severity and a long follow up time up to 14 years. Furthermore, the developed models were well validated in a separate cohort.

Models developed in this study showed a high accuracy to predict five year liver related death, liver decompensation and HCC with an AUROC of 0.95, 0.90 and 0.95 respectively. Cut points were determined to identify patients at higher risk for each clinical outcome and achieved both high sensitivity and specificity. Using the defined cut points, those patients categorised in the high risk group for liver related death, HCC development and liver decompensation had an annual incidence rate of 12.6%, 6.27% and 4.54% respectively. These rates were significantly higher than that of low or moderate risk group. This fulfilled the goal that to use a simple, reproducible and non-invasive method to accurately select those CHC patients who have higher risk of adverse clinical outcomes and thus have greater benefit from antiviral treatment and liver complications surveillance.

Two recent studies followed patients with advanced liver fibrosis for up to five years and tried to develop models that could predict outcomes (218, 219). However, accuracy assessment including AUROC, sensitivity and specificity of these models were not reported in those studies. Another serum model (HCC-4) was developed to predict the risk of HCC development among patients with all
spectrum of liver fibrosis severity and achieved an AUROC of 0.802 (221). Serum fibrosis models including FIB-4, APRI and Lok index were validated using this cohort and achieved an AUROC of 0.87-0.92 for liver related death, 0.75-0.84 for liver decompensation and 0.80-0.87 for HCC development. These results were supported by previous evidences that AUROCs to predict clinical outcomes varied from 0.66 to 0.88 among different fibrosis serum models (211, 216, 254). MELD score showed an inferior predictive ability among all serum models and this further confirmed that MELD score was only useful for short term prognosis for end stage liver disease. Compared with the existing serum models, the LOS panel had advantages in their superior predictive ability and broad utility for CHC patients with all fibrosis severity.

Moreover, specific models were built to predict liver related death, HCC and decompensation and this allowed different serum markers to be included in each predictive model. HA and GGT are well recognised serum markers for clinical outcome prediction and were included in both models to predict liver related death and liver decompensation. Data regarding the use of alcohol and presence of diabetes was not available in the patient cohort studied and therefore the effects of these co-factors on GGT levels could not be assessed. In contrast, ALP and alpha-macroglobulin remained in the final model as independent predictors for HCC development. This suggested that HCC development might proceed through different pathway than the development of liver decompensation.

In summary, this study developed three models (LOS panel) to predict liver related death, liver decompensation and HCC respectively for CHC patients.
The predictive ability of the LOS panel was better than currently used serum models. The LOS panel could potentially improve clinical care by allowing the optimum use of expensive directly-acting antiviral agents before the onset of significant clinical complications. In addition these models would also be valuable in determining the timing of initiation of screening for HCC and for the complications of portal hypertension. Validation studies in a large independent HCV population would be of value however this is unlikely to occur in the short term as there are no other cohorts with long term follow up data readily available. The models also need to be validated in patients with other chronic liver diseases.
8.1. SUMMARY AND DISCUSSION

Chronic hepatitis C is a slowly progressive disease, with an early long term asymptomatic phase. The accurate prediction of fibrosis severity and the prediction of adverse clinical outcomes is a major challenge for individualised management of chronic hepatitis C. While histopathological staging of liver biopsy remains the reference standard for liver biopsy fibrosis assessment a large number of non-invasive methods have emerged in the last two decades as surrogate markers for liver fibrosis assessment and hence clinical outcome prediction. This project has developed and validated a new computerised image analysis method to measure liver fibrosis severity. Novel serum models have also been developed to predict CPA and also to directly predict liver related clinical outcomes.

The strengths of this project was the large well defined cohort that included patients with all stages of liver fibrosis, particularly those with none or minimal fibrosis and the long follow up of up to 20 years. Detailed clinical outcomes for each patient were available for analysis. The whole cohort was included in the first study (chapter 3) to evaluate the long term clinical outcomes for CHC patients stratified by Metavir stage. The results showed that CHC patients with F0, F1 or F2 fibrosis have a benign course with infrequent episodes of liver related morbidity and mortality after 15 years follow up. Those CHC patients
with F3 fibrosis had a similar low rate of liver complications during the first seven years of follow up after which the rate of complications significantly increased. CHC patients with F4 fibrosis had significantly worse survival compared to those with F3 fibrosis. This data suggests a lack of immediate clinical benefit in treating F2 patients as recommended by AASLD and EASL guidelines. The delayed but significant increase in liver complications in those CHC patients with F3 fibrosis suggested that surveillance for HCC and oesophageal varices may need to be extended to those CHC patients with F3 after a follow up period of seven years. Additionally, this study evaluated the ability of Metavir stage to predict adverse clinical outcomes. The results showed that Metavir stage had a poor ability to stratify risk for patients with less severe liver fibrosis.

The next step in this project was to test if image analysis was an improved measurement of liver fibrosis severity compared with Metavir stage. To date non-standardized methods of liver biopsy staining and CPA measurement have greatly limited the clinical use of CPA. To optimise the CPA measurement technology, the study presented in chapter 4 systemically evaluated the influence of the staining method, scanning technique and biopsy size on the reliability and utility of CPA measurement. The results showed that CPA obtained using sirius red stained biopsies had higher accuracy and reproducibility compared to trichrome stained biopsies. CPA values obtained using either 20X or 40X magnification were similar. A biopsy with an analysis area of 5mm² or more could generate reliable CPA values with acceptable sampling variability. Therefore the determined optimum CPA measurement method was to stain a newly cut liver section using sirius red, scan the liver
section at 20X magnification and then calculate the CPA value using pixel counting technique on the digital liver biopsy image. The threshold for positive pixels that corresponded to areas of sirius red staining was determined according to hue value and colour saturation using the original image for comparison. A biopsy with a measurement area less than 5 mm\(^2\) after exclusion of any large portal tracts or capsule was considered as insufficient and thus excluded from further analysis. Compared to other methods, the use of a high resolution scanner with 3.94\(\times\)10\(^6\) pixels /mm\(^2\) and the exclusion of liver capsule and large portal tracts in this study further guaranteed the high accuracy of CPA measurement.

Existing evidence showed that CPA was an independent predictor for HVPG in CHC patients (76), for decompensation in cirrhotic patients (81), for transplantation free survival in children with biliary atresia (249) and for post liver transplantation outcomes in patients with recurrent HCV (82-84). However, the ability of CPA to predict medium to long term clinical outcomes for CHC patients was unknown. The study presented in chapter 5 evaluated the ability of the optimised method of CPA measurement to predict long term clinical outcomes and compared it with that of Metavir stage. CPA values ranged from 1.3\% to 44.6\% among this cohort. This suggested a dramatic increase of CPA value with the progression of liver fibrosis. CPA was well correlated with Metavir stage but considerable overlap of CPA values between each Metavir stage was also observed. This was not surprising as CPA measurement and Metavir stage rely on different aspects of liver fibrosis progression. Metavir stage mainly depends on the severity of architectural distortion of the liver whereas CPA quantitates the amount of liver collagen. CPA values found in cirrhosis (F4)
ranged from 5.9% to 44.6% and the large range suggested that CPA had the potential to provide additional prognostic information in this group of patients.

The continuous nature of the CPA measurement allowed the development of the optimal range of values that stratified incremental risks of adverse clinical outcomes. CPA values were consequently grouped into the following stages: C1: 0%-5% (normal), C2: 5%-10% (minimal), C3: 10%-20% (moderate), C4: ≥20% (severe). CPA stage was able to stratify risk better than Metavir stage with a significant difference in HCC free survival between each consecutive CPA stage. CPA stage remained an independent predictor for liver related death and HCC after adjusting for Metavir stage and age. CPA stage was also significantly associated with the risk of liver related death in cirrhotic patients. These findings support the hypothesis that CPA stage provides improved predictive value compared with Metavir stage.

Another interesting finding of this study was that CPA measurement has the potential to use small biopsies that were insufficient for Metavir staging. Five mm² was used as the minimum acceptable size for CPA analysis and inclusion of these small biopsies resulted in an excellent ability of CPA to predict clinical outcomes. Compared to histopathological staging, CPA measurement has additional requirements that include: a high resolution scanner, image analysis software and around five minute’s analysis time for each biopsy. However, the CPA measurement is semi-automatic, easily performed, requires less expertise and has a higher reproducibility.
As both histopathological staging systems and CPA are based on liver biopsy fibrosis assessment they share similar limitations such as sampling error and risk of serious complications or death. Non-invasive methods such as serum models, elastography (ultrasound-based and MR) and cross sectional radiological imaging have been developed during the last two decades to avoid these limitations. The advantages of serum models compared to radiological tests are their low cost, wide availability and high reproducibility. A large number of serum models have been developed to predict liver fibrosis stage. Some of these models have been widely validated and have a good ability to predict significant fibrosis and cirrhosis. However, all of the currently commercially available serum models were developed using a semi-quantitative histopathological staging system as the reference standard. The well-known limitations of histopathological stages such as inter and intra-observer variability of interpretation and small number of fibrosis stages inevitably decrease the accuracy of these serum models to predict fibrosis severity. This also reduces the ability of serum models to detect fibrosis progression or regression. In contrast, CPA quantitates liver fibrosis with a continuous scale and this allows CPA to have more direct correlation with serum fibrosis markers and this simplifies the interpretation of these values when included in models. As a result, the CPA serum model has the potential to better quantitate fibrosis severity at a single time point and also with sequential assessments. This is of significant clinical importance in the new era of antiviral therapy where fibrosis regression becomes one of the major goals of treatment.

The study presented in chapter 6 developed a serum model (CPAscore) using CPA as the reference standard. CPAscore included HA, α2-macroglobulin and
platelet count and it was closely correlated with actual CPA values. The model fit (R square) of the CPAscore was 0.46, suggesting CPAscore was not the only factor that reflected CPA value. This could be partly explained by the continuous nature of CPA values and serum marker values, both of which may include certain background variations that don’t represent actual fibrosis severity. Individual variation is one important potential variation and it is possible that CPAscore could have higher accuracy to detect fibrosis change for single individual’s by using multiple sequential tests e.g. before and after antiviral treatment. Due to the high cost of antiviral treatment and the lack of a sensitive non-invasive method to detect fibrosis change, CPAscore has the potential to select the most suitable patients for antiviral treatment and to evaluate the efficacy of treatment.

As defined in the previous study, CPA values ≥10% and ≥20% represent the moderate and high risk groups for adverse clinical outcomes. CPAscore achieved good accuracy to detect those patients with CPA larger than 10% and 20% with AUROC of 0.82 and 0.94 respectively. Cut points of 8.7 and 10.7 were used to identify patients with CPA ≥10% and ≥20% respectively and both had high sensitivity and specificity. CPAscore therefore had the ability to stratify patients into low, moderate and high risk group.

In the final study of this project (chapter 7), three serum models [Liver Outcome Score (LOS)] were developed to directly predict liver related survival, HCC and liver decompensation respectively. The potential advantages of serum panels that have been developed to directly predict clinical outcomes are that they will incorporate additional analytes that are not useful to predict fibrosis but are
useful to predict clinical end points. These additional analytes may be associated with other factors such as portal hypertension, coagulopathy, protein synthetic dysfunction and renal failure that are known to predict liver related outcomes. The LOS panels showed a high accuracy to predict five year liver related death, decompensation and HCC with an AUROC of 0.95, 0.90 and 0.95 respectively. Cut points were determined to stratify incremental risk for patients to develop a clinical outcome and they had both a high sensitivity and specificity. Using these cut points patients were categorised into a low, moderate or high risk group for liver related death with the annual incidence rate of 0.06%, 2.57% and 12.6% respectively. Other cut points were used to separate the low and high risk groups for liver decompensation and HCC development with annual incidence rate of 0.20% and 4.54% respectively for liver decompensation and 0.13% and 6.27% respectively for HCC development. This achieved the important goal to develop and validate a simple, reproducible and non-invasive method to accurately select those CHC patients who have higher clinical risk and thus have the greatest potential benefit from antiviral treatment and liver complication surveillance.

In summary, this project determined the optimal method for CPA measurement in liver biopsy samples and it is the largest study that comprehensively evaluated the ability of CPA to predict clinical outcomes. Moreover, four serum models were developed to predict CPA value, the risk of liver related death, HCC development and liver decompensation and were validated in a separate population. These serum models can be used as standard tests in clinical practice and guide individual management of chronic hepatitis C.
8.2. FUTURE DIRECTIONS

8.2.1. Image analysis of liver biopsy

Computerised image analysis of liver biopsy samples has the potential to maximise the information gained from an invasive and potentially dangerous procedure. This project demonstrated the excellent ability of computerised image analysis to measure collagen content of liver biopsies. This technique also has the potential to measure other important histopathological features of liver biopsy.

Image analysis uses a pixel counting technique and this has the potential to measure any molecule that specifically binds a dye. A new threshold would need to be determined to allow the software to recognise the positive pixels that correspond to the stained areas. For example, the area of fat accumulation within hepatocytes can be measured using image analysis of oil red O stained liver tissue. The accumulation of iron within the liver can be measured using Perls Prussian blue stained liver tissue. Moreover, other histopathological features that include nodule size (large, medium, small) and fibrous septa thickness (thick, medium, thin) showed an ability to sub-stage cirrhosis but no attempts were performed to use a computer-based technology to quantitatively measure these features. Image analysis of liver biopsy specimens could potentially calculate the mean value and standard deviation of both the fibrosis septa thickness and nodular area of liver biopsy samples. However, a more sophisticated image analysis program would need to be developed.
8.2.2. CPA measurement

CPA measurement quantitates liver collagen accumulation and can now be used as a prognostic tool for stratifying clinical risk in chronic hepatitis C patients. However this technique needs to be validated in other forms of liver disease and patient populations. A major advantage of the computerised image analysis method developed in this thesis is that CPA can be measured using stored liver biopsy blocks with a high degree of accuracy and reproducibility. Therefore, this allows retrospective studies to be performed using stored tissue, particularly as fewer liver biopsies are being performed on a routine clinical basis. CPA’s ability to accurately detect fibrosis progression or regression would be of great interest. CPA could be used to illustrate the natural history of hepatic fibrosis progression. A number of studies performed liver biopsy on CHC patients who had an identified HCV infection time and CPA would be of potential use in these cohorts to more accurately analyse fibrosis progression. The continuous nature of CPA could allow such studies to answer the long standing question of whether liver fibrosis accumulation in CHC patients has a linear pattern in relationship to time.

Another potential clinical utility of CPA is to sub-stage clinical risk in cirrhotic patients. More studies are needed to evaluate and confirm the ability of CPA to predict HVPG, the presence of esophageal varices, the development of liver decompensation and HCC in cirrhotic patients with all forms of chronic liver diseases.
8.2.3. Serum models

8.2.3.1. CPAscore

This project developed a new serum model (CPAscore) to predict liver fibrosis severity using CPA as the reference standard. Studies that include separate external cohorts are needed to further validate the predictive ability of CPAscore. Moreover, direct serum markers that were not included in this project such as MMP2, TIMP1, N-terminal peptide of procollagen III and laminin have the potential to better correlate with CPA than routine serum tests. Therefore studies that include these candidate serum markers are of interest to test if the predictive ability of CPAscore can be improved by adding additional direct serum markers.

The ability of CPAscore to detect fibrosis progression and regression is an important clinical issue that needs to be addressed using a cohort of patients who have had sequential liver biopsies and serum samples collected, especially before, during and after anti-viral or anti-fibrotic drug treatment.

8.2.3.2. Liver Outcome Score panel

Three serum models (LOS panel) were developed to directly predict liver related death, liver decompensation and HCC respectively. The LOS panel showed excellent correlation with each clinical end point in this project using an Australian population. Validation studies that use external cohorts in United States, Europe and Asian countries are needed to further evaluate their prognostic performance. The predictive ability of LOS panels also needs to be validated in other types of chronic liver disease, such as chronic hepatitis B,
alcoholic liver disease and non-alcoholic fatty liver disease. The combination of additional serum markers specific for each aetiology has the potential to improve the predictive ability of future LOS panels for other types of chronic liver diseases. Furthermore, it will be interesting to evaluate the time-related change of LOS panels and correlate it with clinical outcomes. The changes of LOS panels before and after antiviral therapy can be used as a potential measurement of the efficacy of treatment.
REFERENCES


34. Hellerbrand C. Hepatic stellate cells--the pericytes in the liver. Pflugers Arch 2013;465:775-778.


87. Hui AY, Liew CT, Go MY, Chim AM, Chan HL, Leung NW, Sung JJ. Quantitative assessment of fibrosis in liver biopsies from patients with chronic hepatitis B. Liver Int 2004;24:611-618.


120. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, Lok AS. A simple noninvasive index can predict both


142. Adebajo CO, Talwalkar JA, Poterucha JJ, Kim WR, Charlton MR. Ultrasound-based transient elastography for the detection of hepatic fibrosis in


164. Garcia-Tsao G, Friedman S, Iredale J, Pinzani M. Now there are many (stages) where before there was one: In search of a pathophysiological classification of cirrhosis. Hepatology 2010;51:1445-1449.


