Université de Montréal

Cost-effectiveness of NASH screening

par

Eric W. Zhang, MD

Département des Sciences Biomédicales

Département de radiologie

Faculté de médecine

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de maîtrise en Sciences Biomédicales

Option recherche clinique biomédicale

30, Avril 2017

Université de Montréal

Faculté de médecine

Cost-effectiveness of NASH screening

par

Eric W. Zhang, MD

a été évalué par :

Dr. Catherine Beauchemin : Président rapporteur

Dr. An Tang : Directeur de recherche

Dr. Luigi Lepanto : Co-directeur de recherche

Dr. Nandini Dendukuri : Membre du Jury

Résumé

La stéatose hépatique non alcoolique (NAFLD) est la cause d'hépatopathie chronique la plus fréquente dans les pays occidentaux. Aucune étude n'a vérifié le rapport coût-efficacité du dépistage pour la stéatohépatite non-alcoolique (NASH), le stade avancé de la maladie.

Nous avons réalisé une analyse coût-utilité des stratégies annuelles non invasives de dépistage, utilisant une perspective d'un système de soins canadien, dans la population générale et l'avons comparé à une population à haut risque composée de patients obèses et diabétiques. Les algorithmes de dépistage incluent des techniques bien étudiées notamment le «NAFLD fibrosis score», la technique «transient elastography» (TE), et l'imagerie «acoustic radiation force impulse» (ARFI) pour la détection de la fibrose avancée (≥ F3); et le test «plasma cytokeratin-18» (CK-18) pour la détection de la NASH. La biopsie du foie et l'élastographie par résonance magnétique (MRE) ont été comparées comme méthodes de confirmation. Les coûts en dollars canadiens furent corrigés en fonction de l'inflation et actualisés à un taux d'actualisation de 5%. Un rapport coût-efficacité différentiel (ICER) de ≤$C50,000 / année de vie pondérée par la qualité (QALY) a été considéré comme coût-efficace.

Nous avons trouvé que par rapport à la stratégie sans dépistage annuel, le dépistage annuel avec l'algorithme NAFLD fibrosis score/TE/CK-18 et avec MRE comme méthode de confirmation pour la fibrose avancée, a donné un ICER de $C26,143 par année de vie pondérée par la qualité (QALY) gagnée. Le dépistage annuel dans les populations à haut risque obèses et diabétiques était encore plus coût-efficace, avec un ICER de $C9,051 et $C7,991 par QALY gagnée respectivement. La confirmation avec la biopsie du foie n'était pas coût-efficace.

Notre modèle indique que le dépistage annuel pour la NASH peut être coût-efficace, particulièrement dans les populations obèses et diabétiques à haut risque.

Mots clés: Stéatose hépatique non alcoolique (NAFLD), stéatohépatite non-alcoolique (NASH), coût-utilité, dépistage, coût-efficacité, élastographie

Abstract

Nonalcoholic fatty liver disease (NAFLD) is the most common liver disease in Western countries. No studies have examined the cost-effectiveness of screening for nonalcoholic steatohepatitis (NASH), its advanced form.

We performed a cost-utility analysis of annual non-invasive screening strategies using third-party payer perspective in a general population and compared it to screening in a high-risk obese or diabetic population. Screening algorithms involved well-studied techniques including NAFLD fibrosis score, transient elastography (TE), and acoustic radiation force impulse (ARFI) imaging for detecting advanced fibrosis (≥ F3); and plasma cytokeratin-18 for NASH detection. Liver biopsy and magnetic resonance elastography (MRE) were compared as confirmation methods. Canadian dollar costs were adjusted for inflation and discounted at a 5% rate. Incremental cost-effectiveness ratio (ICER) of ≤$C50,000 / quality adjusted life year (QALY) was considered cost-effective.

Compared with no screening, screening with NAFLD fibrosis score/TE/CK-18 algorithm with MRE as confirmation for advanced fibrosis had an ICER of $C26,143 per quality-adjusted life year (QALY) gained. Screening in high-risk obese or diabetic populations was more cost-effective, with an ICER of $C9,051 and $C7,991 per QALY gained respectively. Liver biopsy confirmation was not found to be cost-effective.

Our model suggests that annual NASH screening can be cost-effective, particularly in high-risk obese or diabetic populations.

Keywords: Nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), cost-utility, screening, cost-effectiveness, elastography

Table of Contents

List of Figures

List of Tables

Abbreviations

AASLD	American Association for the Study of Liver Diseases
ACG	American College of Gastroenterology
AFLD	Alcoholic fatty liver disease
AGA	American Gastroenterological Association
ALT	Alanine aminotransferase
ARFI	Acoustic radiation force impulse
AST	Aspartate aminotransferase
AUROC	Area under a receiver operating characteristic curve
BMI	Body Mass Index
CK-18	Cytokeratin-18
CVD	Cardiovascular disease
DPCP	Detectable preclinical phase
EASL	European Society for the Study of the Liver
ELF panel	Enhanced liver fibrosis panel
HCC	Hepatocellular carcinoma
HCV	Hepatitis C
ICER	Incremental cost-effectiveness ratio
MRE	Magnetic resonance elastography
NAFL	Nonalcoholic fatty liver
NAFLD	Nonalcoholic fatty liver disease

NASH	Nonalcoholic steatohepatitis
QALY	Quality-adjusted life year
TE	Transient elastography

Acknowledgements

I express my sincere gratitude to my principal supervisor, Dr. An Tang, who first introduced to clinical research and helped inspire the basis behind this thesis. He is the model for scientific rigor, methodology, and an expert scientist, whose research in hepatic imaging has contributed to the adoption of imaging biomarkers in chronic liver disease. I consider myself very fortunate to have Dr. Tang as a supervisor and mentor over the years. His boundless enthusiasm, exemplary professionalism, and admirable creativity have inspired me and will continue to illuminate the path in my future life and career.

Dr. Luigi Lepanto, I thank for his continued and unwavering commitment and support. His expertise in Health Economics and clinical wisdom was invaluable throughout my Master's degree and for the completion of my thesis. He is an excellent scholar, physician, and researcher. I am grateful for your continued support.

My gratitude as well to my professors and colleagues at the LBUM and University of Montreal, including Dr. Siavash Kazemirad, Dr. Jean Lachaine, and Dr. Guy Cloutier, whom I have had the opportunity and the pleasure to work with. Your support was invaluable as well to the completion of my studies.

Last, but certainly not least, I would like to also thank my family, who have been unwavering in their support during the completion of my Master's degree. Their encouragements have allowed me to realize my projects and my goals.

Chapter 1

Nonalcoholic fatty liver and the need for health economic assessment

1.1 Introduction

Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease in the Western hemisphere. It encompasses a spectrum of disease arising from the accumulation of fat in the liver. Over time, the fat depositions cause inflammatory changes within the liver, leading to non-alcoholic steatohepatitis (NASH), a subgroup of NAFLD. In turn, chronic inflammation of the liver results in collagen deposition and liver fibrosis. This process continues until the hepatic parenchyma is irreversibly changed and end-stage cirrhosis is reached. Cirrhosis is associated with numerous important clinical implications, including end-stage liver failure and increased risk of hepatocellular carcinoma. (1)

Morphologically, NAFLD is indistinguishable from alcoholic fatty liver disease. However, as the name would imply, in NAFLD, the deposition of liver fat occurs in the setting where there has been no significant alcohol consumption. (2) Instead, its pathophysiology and genesis relates to the amalgam of risk factors known collectively as metabolic syndrome. (3) It is by no coincidence that the rise in prevalence of NAFLD over the last two decades has paralleled the equally significant increases in the prevalence of obesity and type two diabetes. (4)

Currently, there is controversy regarding whether or not to screen for NAFLD/NASH. The European Society for the Study of the Liver (EASL) recommends screening for the disease in high-risk patients with metabolic syndrome and/or patients with characterized insulin resistance. (1) On the other hand, the American Gastroenterological Association (AGA), American Association for the Study of Liver Diseases (AASLD), and American College of Gastroenterology (ACG) do not recommend screening, even in high-risk groups with obesity or type two diabetes. (5)

These differing opinions stem from uncertainty in current literature regarding diagnostic tests, treatments, and overall healthcare cost-effectiveness of screening for NAFLD/NASH. In accordance with the classic screening criteria established by Wilson and Jungner, there should be an agreed-upon policy for diagnosis and treatment. In particular, the cost of such a screening policy should be balanced economically with medical expenditure as a whole.

(6) In addressing this important point, this thesis aims to address the current knowledge gap regarding the cost-effectiveness of screening for NAFLD/NASH.

At present, the gold standard for the diagnosis of NAFLD and NASH is liver biopsy. However, biopsy is not a benign test and entails significant risks of morbidity and mortality. These risks include bleeding, infection, and on rare occasions, even death. Combined with other limitations such as cost and sampling error, liver biopsy would be an unacceptable test for screening purposes. (7, 8) Thus, to address this issue, noninvasive detection methods have emerged over the last few decades. Among these, the most numerous and widely used noninvasive tests include: serum markers and elastography techniques. Individually, these tests have many strengths including the obvious noninvasive nature of the exam, the relative low cost, and relative ease of use. Some screening serum markers may be limited by characteristic weaknesses such as lower receiver operating characteristic (AUROC) performance ratings for detecting disease. Nonetheless, when combined together in an algorithm with both screening serum markers and diagnostic elastography techniques show promise in the screening of NAFLD. (9)

Ultimately, the goal of any screening program should be to recognize latent or early symptomatic stage of disease in hopes of diagnosing and treating the illness before the disease fully declares itself in terms of morbidity and mortality. In the case of NAFLD and NASH, the goal would be to detect and treat early stages of disease before patients reach irreversible liver cirrhosis and its associated costly sequela. While the exact cost of NAFLD/NASH to the healthcare system is difficult to quantify, it is estimated to be substantial. (10) More than a decade ago, chronic liver disease and liver cancer accounted for approximately 3 billion dollars in American healthcare cost. Furthermore, it was the 10^{th} most common cause of disease-related death in the United States. (11) This number is now estimated to be much higher, given the increasing trend in metabolic disorder, obesity and type two diabetes. NAFLD/NASH is set to become the leading cause of liver transplantation by 2020. (12)

1.2 Most prevalent chronic liver disease

Liver disease is an important cause of morbidity and mortality in Western society and around the world. (13) In the late 1990s in the United States, it was estimated to account for 2% of all deaths and 1% of all health care expenditures. (11) Recently, the Centers for Disease Control and Prevention found that liver disease is the 12th leading cause of all death in the United States. (13) A population-based study from the United States demonstrated that the prevalence of chronic liver disease has climbed significantly from 11.78% in the years 1988-1994 to 14.78% in the years 2005-2008. (14) The increasing healthcare burden of liver disease is likely related to and exacerbated by the increasing prevalence of obesity and type 2 diabetes mellitus, which are known risk factors for NAFLD/NASH. Presently, NAFLD represents the most common cause of liver dysfunction. (4) Currently, more than two-thirds of the American population are either overweight or obese. (15) These alarming trends are similar in other Western countries, including Canada. (16) **Figure 1.1** demonstrates the rise in prevalence of obesity in Canada over the last 30 years.

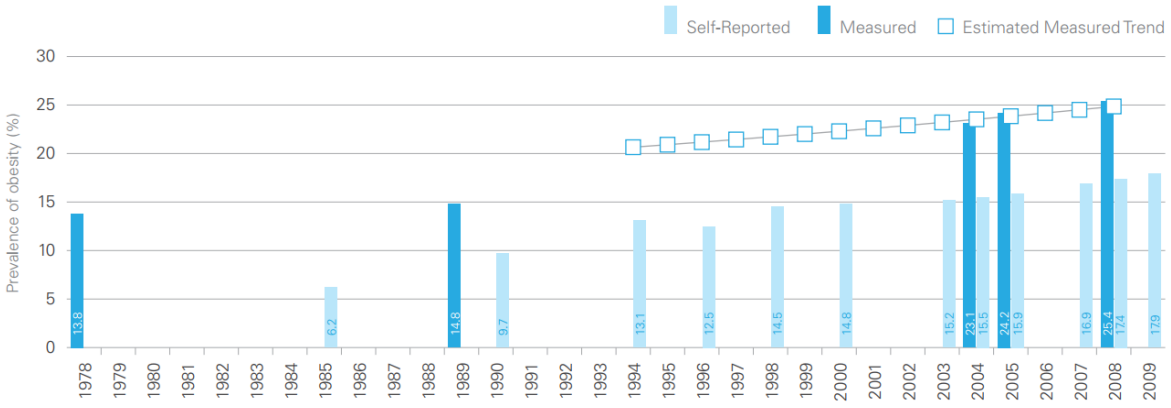


Figure 1.1: Prevalence of Obesity, Ages 18 Years and Older, Canada, 1978-2009. Obesity in Canada [Internet]. Public Health Agency of Canada [cited 2014 Feb 14]. Available from: http://www.phac-aspc.gc.ca/hp-ps/hl-mvs/oic-oac/assets/pdf/oic-oac-eng.pdf.

While liver disease as a whole has been on the rise, it is now recognized that NAFLD is becoming an important cause of chronic liver disease. Current literature estimates that some 280 million obese individuals are affected by NAFLD. (17) Over the last three decades, the prevalence rates for other traditional leading causes of chronic liver disease, such as hepatitis C (HCV) and alcoholic fatty liver disease (AFLD), have not changed significantly over this

time. On the other hand, NAFLD has steadily risen in prevalence every year. The prevalence of NAFLD in the United States rose from 5.51% between the years 1988-1994, to 9.84% between the years 1999-2004, to 11.01% between the years 2005-2008. (14) Looking at the indications for liver transplantation in end-stage liver disease, Charlton et al. found that NASH is the third most common reason for transplantation. The trend showed that NASH as a reason for transplantation increased every year from 2001 to 2009 from 1.2% to 9.7%. On the other hand, the current number one and two reasons for liver transplantation, HCV and AFLD, have been trending downwards each year. Based on statistical projections, NASH would become the leading cause of liver transplantation between 2020 and 2030. (12, 18) **Figure 1.2** demonstrates the projected trend for liver transplantation indication.

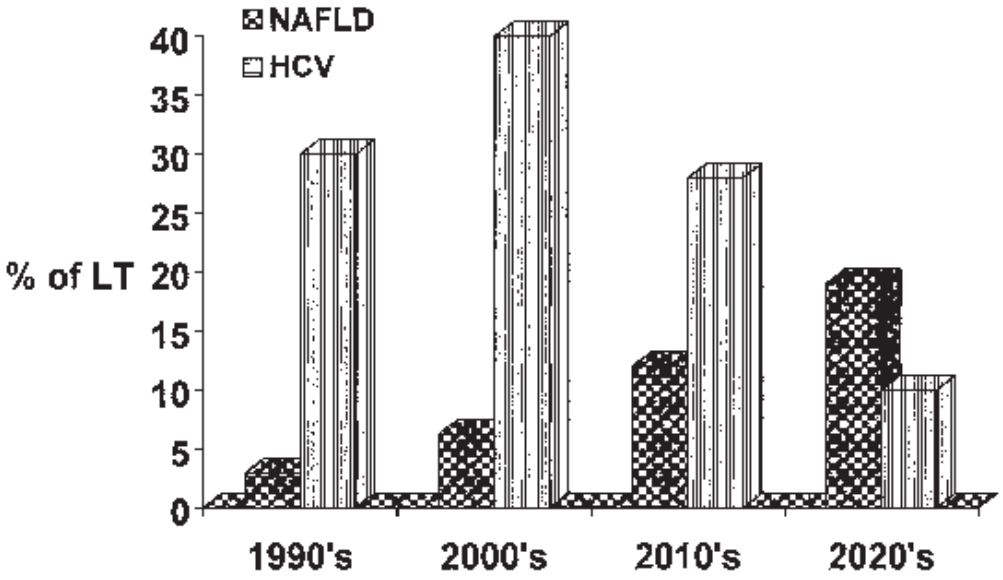


Figure 1.2: Projected relative frequencies of NASH and HCV as indications for liver transplantation. Charlton M. Nonalcoholic fatty liver disease: a review of current understanding and future impact. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2004;2(12):1048-58.

1.3 Global trends in nonalcoholic fatty liver disease

The predominance of NAFLD in Western countries has been well established. However, it is rapidly becoming apparent that NAFLD is not solely a Western phenomenon, but rather that it is becoming a global epidemic. (19) With increased globalization, job market

modernization, and urbanization of the populace, emerging economies have given way to the adoption of a more sedentary lifestyle, leading to a rise in the prevalence of NAFLD.

For example, several studies have already found that prevalence rates in China are now comparable to those in Europe, at 15-30% of the general population. (20, 21) Similarly in Indian urban populations, the prevalence of NAFLD ranges from 16 to 32%. Interestingly, in rural areas of the Indian subcontinent, the prevalence is only ~9%, which gives weight to the association between the modern, sedentary lifestyle and increased NAFLD prevalence. (22-24) There is a paucity of epidemiological studies on NAFLD in Africa. However, prevalence rates in Latin America and Australia have been found to be similar to those in North America and Europe. (25, 26) Based on these studies, it is estimated that some one billion people around the world are currently affected by NAFLD. (19) In the coming years, as more and more people are lifted from rural poverty, the prevalence of NAFLD is expected to continue its significant rise around the globe. **Table 1.1** summarizes the high prevalence values of NAFLD stratified by region globally.

Table 1.1: Prevalence of NAFLD globally

Region	**Prevalence (%)**	**95 CI (%)**
Africa	13.48	(5.69-28.69)
Asia	27.37	(23.29-31.88)
Europe	23.71	(16.12-33.45)
Middle East	31.79	(13.48-58.23)
North America	24.13	(19.73-29.15)
South America	30.45	(22.74-39.44)

Adapted from: Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64(1):73-84

1.4 Economic health assessment and decision analysis

A large part of implementing new healthcare strategies revolves around the concept of economic analysis. It is often not enough to establish the efficacy of a drug or a screening

strategy, but also address the question of effectiveness and cost-effectiveness. Put more plainly, it should not only work on an individual level, but also be worth it when implemented into society. In the current climate of growing healthcare cost, the notion of responsible healthcare expenditure is especially important for health policy makers and healthcare professionals alike. (27) It is therefore necessary to address the cost-effectiveness of any new healthcare screening program. (28)

Decision analysis trees are used to quantify and compare various healthcare strategies, and have been largely established in the field of pharmacoeconomics. It is an elegant way of simplifying complex strategies into different decision nodes, and quantifies the differences and consequences associated with each decision. (29) **Figure 1.3** is an example of a standard decision tree for deciding between taking a drug and not taking a drug for a generic disease.

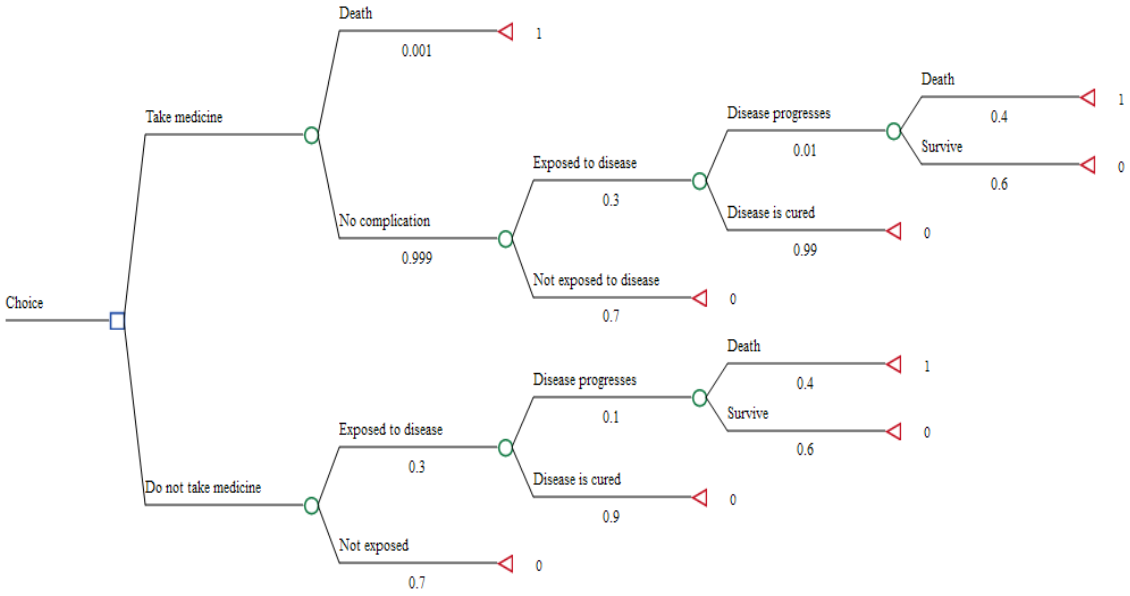


Figure 1.3: Example of a standard decision tree. The blue square denotes a choice node between competing options. The green circle denotes a probability node, indicating the probability of an event occuring. The red triangle denotes a termination node, indicating the end of a decision branch. The probability values are indicated under each event branch. Each termination node is assigned an outcome variable. In this case, the outcome of interest is death, which is assigned a "1". Thus, in this case, we are calculating the difference in deaths between taking and not taking the medicine.

A Markov model is a stochastic model with three key features which make them particularly useful in addressing chronic diseases and clinical scenarios. Firstly, the simulated population begin in a finite set of mutually exclusive health states. Secondly, there is an established time period, called a “cycle”. Each cycle, individuals either move onto another health state or stay in their current health state. Thirdly, movements between health states each cycle is governed by a transitional probability. The main advantage of Markov modeling in decision analysis is that it allows simulations that are more complex, and therefore, more in line with real life. More possible events can be simulated and over a longer time period. (30)

Chapter 2

Nonalcoholic fatty liver disease: an overview

2.1 Definition

NAFLD is a broad spectrum of disease characterized by excessive fat accumulation (>5% macrovesicular steatosis) arising in a setting where there is no significant alcohol consumption and where other causes of liver disease (viral, genetic, autoimmune etc.) have been excluded. The American societies define significant alcohol consumption as > 21 drinks on average per week in men and > 14 drinks on average per week in women. (5)

NAFLD has two major subdivisions: nonalcoholic fatty liver (NAFL or simple steatosis) and NASH (nonalcoholic steatohepatitis). Simple steatosis is the non-progressive form of NAFLD that rarely develops into NASH and more serious sequelae of chronic liver disease. NASH is defined as the subgroup of NAFLD characterized by the presence of steatosis, ballooning degeneration and lobular inflammation, with or without peri-sinusoidal fibrosis. (31, 32) NASH is the progressive form of NAFLD that can advance to fibrosis, cirrhosis, hepatocellular carcinoma (HCC) and liver-related mortality. (33) **Figure 2.1** illustrates the NAFLD disease continuum.

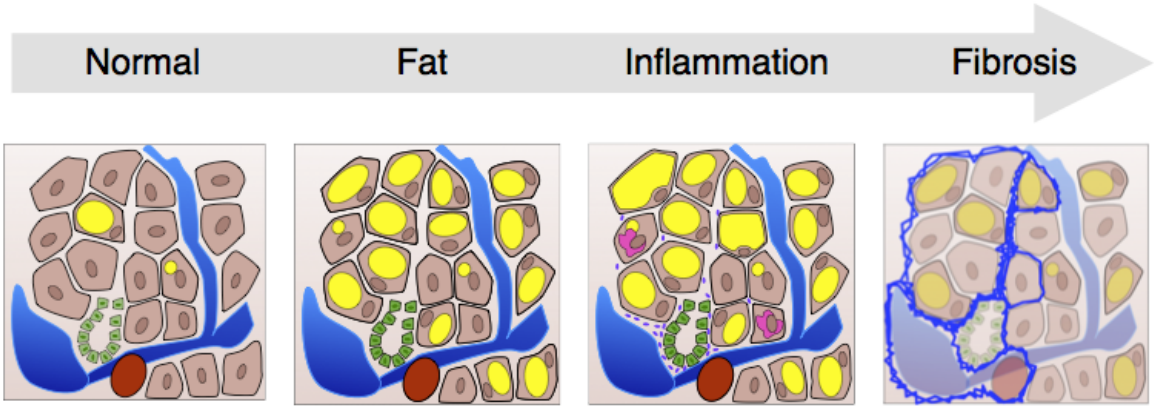


Figure 2.1 (*courtesy of Dr. An Tang*): NAFLD disease continuum. In nonalcoholic fatty liver or simple steatosis, there is >5% fat infiltration with or without mild inflammation (denoted intracellularly in yellow). As the disease progresses, so does the necro-inflammatory changes including ballooning degeneration, Mallory bodies (denoted intracellularly in pink), and inflammatory cell infiltration of the liver (denoted by the tiny purple cells). Chronic inflammation leads to increasing liver fibrosis (denoted in blue).

2.2 Epidemiology

2.2.1 NAFLD in the general population

Population-based studies on NAFLD prevalence have been done using a variety of diagnostic methods. While it is probably most accurate to use liver biopsy, the reference standard for diagnosis, the invasive nature of the exam makes it unsuitable for epidemiological studies. Somewhat circumventing this issue, an American study looking at liver biopsies of potential liver donors found that 20% of potential donors were ineligible for organ donation due to significant degrees of steatosis (>30%). (34) Elsewhere in South Korea, among more than 500 consecutive potential liver donors, the prevalence of NAFLD was even higher at 51%. (35)

Given the risks behind using liver biopsy in epidemiological studies, most studies looking at the prevalence of NAFLD have used non-invasive methods such as imaging-based studies and serum markers. Evidently, these non-invasive methods are less accurate than histology-established diagnosis. Nonetheless, significant conclusions can be drawn from these large population-based studies.

Non-invasive imaging-based studies have assessed the prevalence for NAFLD using magnetic resonance imaging (MRI) and ultrasonography. A large multicenter, cross-sectional population study in Spain demonstrated that the prevalence of NAFLD was 33.4% in men and 20.3% in women. (36) In the Dionysos nutrition and liver study in Italy, subjects with and without suspected liver disease underwent ultrasonographic testing for NAFLD. The study found that among 3,345 subjects, the prevalence of NAFLD in those with and without suspected liver disease was 25% and 20% respectively.(37) American studies have found that the prevalence of NAFLD is equally high if not higher. A large ultrasound-based study performed at the Brooke Army Medical Center found the prevalence of NAFLD to be 46%. (38) The Dallas Heart Study, which used a MRI-based (MR spectroscopy) method for detecting NAFLD, demonstrated that the general prevalence of NAFLD in the general population was 31%. (39)

In terms of serum markers, alanine aminotransferase (ALT) and aspartate aminotransferase (AST) have been used as non-invasive indicators of NAFLD in population-based studies. A study by the Johns Hopkins Hospital Execute Health Program demonstrated

that 14% of NAFLD patients had at least one elevated liver enzyme while 21% had both.(40) Another study estimated the prevalence of NAFLD using only aminotransferases to be around 8% to 9%.(41) However, it is worth noting that while aminotransferases is relatively cheap and readily available, it is not a great test for diagnosing NAFLD. It is now known that a significant number of patients with NAFLD have normal ALT and AST levels.(33) In the Dallas Heart study, 80% of cases of patients with increased hepatic triglyceride content were reported to have normal aminotransferases. (37) Due to its low specificity, aminotransferases alone is unlikely to provide an accurate assessment of population-wide prevalence.

In summary, most American studies have reported the prevalence of NAFLD to be 10-35%. Rates reported from the rest of the world is surprisingly similar, ranging from 6% to 35%, with a median of 20%. (33) The discrepancy between the rates likely varies with the modality used as well as the study population. In general, the prevalence of NAFLD in the North American is thought to be closer to 30%, given that approximately one-third of the population is obese. (33)

2.2.2 Risk factors for NAFLD

Epidemiological studies looking into NAFLD risk factors reveal that excessive BMI and visceral obesity are among the most important risk factors. Bariatric surgery patients with severe obesity have been found to have a prevalence of NAFLD exceeding 90%. Up to 5% of these patients have undetected NASH cirrhosis as well. (33, 42, 43) The other major risk factor for NAFLD is type two diabetes mellitus. Studies have shown that between 69 to 87% of type two diabetic patients assessed by either ultrasound or biopsy demonstrated some form of NAFLD. (44, 45) Other risk factors include individuals with dyslipidemia, of which the prevalence of NAFLD is estimated to be 27-92%. (4) Taken as a whole, these major risk factors for NAFLD are essentially those described in metabolic syndrome (MS), with the two entities being intimately associated. **Figure 2.2** illustrates the average prevalence of NAFLD and NASH in obese and diabetic patients.

Besides metabolic syndrome, other factors associated with NAFLD include age, male gender, and Hispanic heritage. The prevalence of NAFLD increases with age. Advanced age is also linked to increased likelihood of developing NASH cirrhosis and mortality risk

associated with NAFLD. (4) In terms of gender, males are considered more at risk of developing NAFL. (46) Finally, the Hispanic population have been found to have significant higher prevalences of NAFLD than non-Hispanics. (47)

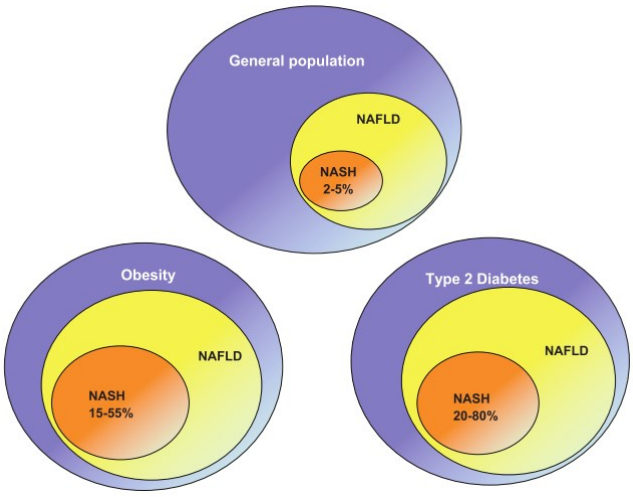


Figure 2.2: The average prevalence of NAFLD and NASH in general and high-risk groups. Bhala N, Jouness RI, Bugianesi E. Epidemiology and natural history of patients with NAFLD. *Curr Pharm Des*. 2013;19(29):5169-76.

2.2.3 NAFLD in the pediatric population

NAFLD in children is a significant known entity that should be recognized early to offset the rapid development of severe complications. Children as young as 2 years old have been found to have NAFLD, with documented cases of NASH-related cirrhosis as young as 8 years old.(48, 49) In such extreme cases, genetic or environmental susceptibility may be called into question.

Given that the definition for NAFLD is the same for children as adults, precise estimation of prevalence in this subset population presents with the same difficulties as in adults. The different estimates vary depending on the type of imaging or serum test, the cut-offs for detection, as well as geographic differences in age, sex and ethnicity. An autopsy study using the gold standard liver biopsy estimates NAFLD prevalence to be 9.6% in 742 children aged 2 to 19 years old who died from unnatural causes.(48) Another study using abnormal aminotransferases as serum detection cites the prevalence of NAFLD in 17-18 year olds to

be 23%.(49) Multivariate analyses have demonstrated that obesity, male gender and older age are independent predictors of fatty liver prevalence in children. (49)

2.2.4 NASH cirrhosis and liver transplantation

With the rise of NAFLD, NASH cirrhosis is expectedly becoming an increasingly common reason for orthotopic liver transplantation. According to the United Network for Organ Sharing, NASH cirrhosis accounted for 3.5% of transplants in 2005, versus just 0.1% of transplants in 1996. Furthermore, the number of transplantations attributed to NASH cirrhosis is thought to be significantly higher, due to its under-recognition as well as association with exclusive comorbidities such as obesity or diabetes mellitus. (50) The number of patients who will undergo transplantation due to NASH cirrhosis is expected to rise in the coming decade as recognition of NAFLD improves, the obesity epidemic worsens and the prevalence of hepatitis continues to decrease.(12)

Interestingly, about 25% of patients with transplanted liver for NASH cirrhosis redevelop steatosis in the first year. By the fourth year, almost 50% of patients will develop steatosis, with 30-50% of these patients also demonstrating histologic evidence of NASH. (51) Post-transplantation studies have found that risk factors for recurrent or de novo NAFLD include obesity, diabetes mellitus/insulin resistance and elevated total cholesterol, the very same risk factors for the development of NAFLD in the first place. (50)

In summary, NASH is becoming an increasingly important reason for end-stage liver disease. It has surpassed alcoholic liver disease, falling behind only hepatitis C as the second-leading indication for liver transplantation. (52)

2.3 Natural history

The actual trigger for the evolution of simple fatty liver disease to end-stage fibrosis/cirrhosis is not well understood. Current literature on this subject is lacking in terms of well-controlled, longitudinal studies, and is limited by use of nonstandard definitions as well as referral and publication bias. Null studies concerning disease progression are less likely to be submitted and published. Furthermore, well designed longitudinal studies are more time-consuming and costly. (51) Of all potential predictors of

disease progression, initial patient histology on presentation has demonstrated the best predictive value for progression of disease. While those with "benign" fatty liver appears to have a small likelihood to progress to cirrhosis over a single lifetime, it is those patients with inflammation or histopathologically-proven steatohepatitis that have the increased likelihood of advancing to fibrosis. (53) Importantly, it is patients with NASH and advanced fibrosis that have the greatest risk of developing cirrhosis, liver failure, and hepatocellular carcinoma. (54, 55) Patients with isolated steatosis demonstrate very low progression to fibrosis and liver-related mortality. (56) Patients with biopsy-proven NASH NAFLD are six times more likely to die from liver-related mortality than non-NASH NAFLD. (57)

On the other hand, patients with NASH are at increased risk of developing early and advanced fibrosis. Current literature places fibrosis progression in NASH at 25 to 30% of cases over 4 years and in 50% of cases over 6 years.(31, 56, 58-61) Other estimates place the rate of progression at one fibrosis stage every 7 years, which is significantly higher than rates seen in non-NASH NAFLD. (62) Once progressive fibrosis begins, the patients are at risk of developing end-stage liver cirrhosis. According to one of the longest cohort studies on this subject, over a mean period of 13.7 years, 13% of patients with mild-to-moderate fibrosis developed cirrhosis. Furthermore, 25% of patients with moderate-advanced fibrosis developed cirrhosis and end-stage liver disease. (56) **Figure 2.3** illustrates the Kaplan-Meier survival curve demonstrating the difference in progression to advanced fibrosis (stage 3 or 4 fibrosis) between patients with biopsy-proven inflammation and those without.

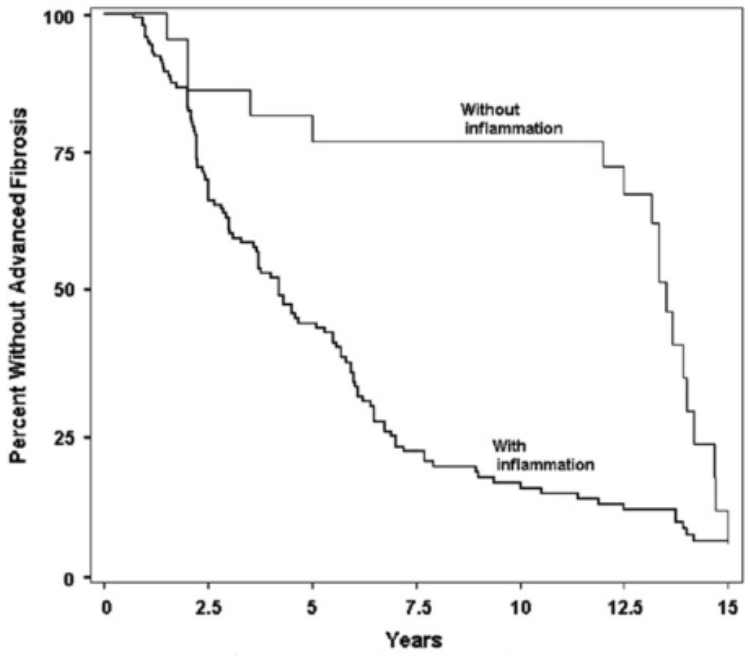


Figure 2.3: Kaplan-Meier survival curve for the difference in progression to advanced fibrosis between patients with inflammation and those without. Advanced fibrosis is defined as stage 3 or stage 4 fibrosis. The two cohorts are stratified by the presence of any inflammation on initial index biopsy in patients included in paired biopsy, natural history studies of NASH. Argo CK, Northup PG, Al-Osaimi AM, Caldwell SH. *Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis*. Journal of hepatology. 2009;51(2):371-9.

Metabolic syndrome plays a central role in the development of NAFLD. Metabolic syndrome consists of central obesity, hypertension, hypertriglyceridemia, impared glucose tolerance and low high-density lipoprotein cholesterol. Over 90% of NAFLD patients have at least one of these aforementioned traits.(63) The acceleration of NAFLD over the last two decades has paralleled the rise of overweightness and obesity in the general population. In the Dionysos study, 94% of obese patients (body mass index greater than or equal to 30kg/m^2) and 67% of overweight patients (body mass index greater than or equal to 25/m^2) had NAFLD. These staggering figures were compared with only 25% of normal weight patients who had NAFLD. (37, 64) Abdominal obesity appears to be an independent predictor for NAFLD, with an association noted between degree of abdominal obesity and the likelihood of NAFLD. (65)

The other important factor in the pathophysiology of NAFLD is insulin resistance. The overall prevalence of NAFLD in those patients with type 2 diabetes is between 40% and 70%. (66) Many cross-sectional studies have postulated that hepatocyte ballooning, inflammation and fibrosis, phenotype changes in steatohepatitis are associated with type 2 diabetes. (63) There is also evidence of higher degrees of steatosis in type 2 diabetes with up to 200% more liver fat than those without in age-, gender-, and BMI-matched controls. (67)

Interestingly, according to one of the largest multicenter prospective study from four countries, end-stage cirrhosis from NAFLD appears to lead to lower rates of liver-related complications as well as lower rates of HCC as compared with end-stage cirrhosis by HCV infection. However, the overall mortality of both conditions is very similar, due in part to similarities in vascular/non-liver related complications. This important fact harkens back to the fact that hypercholesterolemia and diabetes are associated with both NAFLD as well as major vascular complications. Thus, the authors argue for more holistic treatments for NAFLD in order to tackle the very real risk of mortality from cardiovascular complications. (68)

In summary, the presence of inflammation or NASH on initial liver biopsy comprises one of the strongest predictors for NAFLD progression. The degree of fibrosis is the most important prognostic factor. (51, 55) Liver steatosis by itself is not associated with increased liver-related mortality or significant NAFLD progression. Given this knowledge on the natural history, the effort should be oriented towards identifying and treating the patients with steatohepatitis and advanced fibrosis.

2.3.1 Prognostic differences between NAFLD histological subtypes

A meta-analysis of five community-based studies assessing the prognosis of different NAFLD histological subtypes demonstrated interesting results.(9) The survival of those patients with simple steatosis was found to be very similar to that of the general population. In comparison, the patients with NASH had a significantly higher overall mortality than those with simple steatosis with an OR of 1.81 (CI 95%). Of all possible causes of death, liver

disease was the main etiology for death excess in NASH, with a liver-related mortality rate of 11 to 17.5% as compared with 1.7 to 2.7% in patients with simple steatosis. (9)

While the presence of NASH is a significant factor in liver-related and overall mortality, fibrosis stage is an even stronger predictor of overall mortality.(55) Besides overall mortality, fibrosis stage is also a strong predictor for hepatocellular carcinoma and cirrhosis, two long-term liver-related complications of NAFLD. Interestingly, fibrosis stage has also been found to be predict increased rate of cardiovascular and infectious diseases. (55)

2.3.2 Liver-related complications

Liver-related complications in NAFLD can be divided into those related to cirrhosis and terminal liver failure, and long-term complications related to hepatocellular carcinoma (HCC). Cirrhosis imparts increased mortality and complication risk, related to synthetic liver dysfunction, which is a well-studied process. (4) In theory, patients with advanced NAFLD who reach the cirrhotic stage would suffer from the same rate of complications as cirrhotic patients from any other etiology such as hepatitis C (HCV). However, there is evidence to suggest that NAFLD cirrhosis has lower rates of liver-related complications when compared with HCV patients. A large multicenter prospective trial of 511 patients comparing the mortality and morbidity between NAFLD and HCV cirrhotic patients found that the cumulative incidence of liver-related complications was lower in the NAFLD cohort than the HCV cohort. (68) Importantly, the same study did not demonstrate any significant difference in overall mortality between the two cohorts. (68) These results were confirmed in previous smaller trials as well.(69, 70)

HCC represents another significant complication associated with advanced NAFLD and cirrhosis. While population-based studies looking at the long-term history of NAFLD have been limited by small size, the bulk of the evidence have confirmed the association between NAFLD and HCC.(31, 68, 70) One of the largest prospective community-based study thus far demonstrated a 10% rate of HCC in patients with NAFLD cirrhosis after a mean follow up of 7.6 years.(71) Interestingly, it has been demonstrated that there is a relative decrease in risk of HCC in patients with NASH cirrhosis versus HCV cirrhosis. A large cohort study

concluded that the risk of HCC in HCV was significantly more than in NAFLD (6.8% versus 2.4% respectively).(68) Nonetheless, HCC remains an important complication of NAFLD. NAFLD cirrhosis is associated with an incidence of up to 10% over 7 years.(4) Furthermore, there has been evidence to suggest that HCC can develop in NAFLD patients even in the absence of cirrhosis, where metabolic syndrome is the only identifiable risk factor.(72) Lastly, most cases of HCC associated with NAFLD are detected on first referral, a fact that outlines the importance of clinical vigilance and surveillance of disease.(4)

2.3.3 Extra-hepatic complications

Besides liver-related complications, NAFLD also increases the risk for developing type 2 diabetes, cardiovascular disease (CVD), and peripheral vascular disease. The relationship between NAFLD and risk of future diabetes was established in a meta-analysis of 21 prospective, population-based studies.(73) Furthermore, patients with both NAFLD and type 2 diabetes have a further increased risk of diabetes-related complications such as coronary heart disease, ischemic stroke, and cardiovascular death.(74)

Growing literature seems to suggest an independent and active involvement of NAFLD in the pathogenesis of CVD.(75) It has been demonstrated that the presence of hepatic steatosis is associated with increased intima-media thickness of the carotid arteries as well as increased presence of carotid plaques.(76) Furthermore, patients with NAFLD have been associated with significantly higher estimated cumulative risk of major cardiovascular events (19% in NAFLD patients vs. 10% control).(77) However, while the association exist, the underlying causal mechanism linking the two pathologies is currently uncertain.

2.4 Screening, surveillance, and diagnostic modalities

Given the previously mentioned uncertainties, the specifics of screening and surveillance policies in NAFLD remain an open question.(78) Nonetheless, much work is currently underway in finding the most accurate noninvasive modalities in the diagnosis of NAFLD

across its disease spectrum. These noninvasive modalities are divided into two broad categories, biomarker panels and imaging-based methods.

2.4.1 Biomarker panels

Biomarker panels represent a significant proportion of the current tools available for the non-invasive assessment of NAFLD. They include an array of biochemical parameters ranging from routine liver function tests to markers of hepatocyte apoptosis and markers of adipose tissue-releasing cytokines. Despite universal reliance on liver transaminases for the detection of liver pathology, in reality liver transaminases by themselves are not accurate enough for NAFLD screening. This is because the majority of patients with NAFLD present with normal transaminase levels and histologically advanced disease can readily be missed.(78) While individually, blood tests may lack in diagnostic accuracy, when combined together in an algorithm, their diagnostic accuracy increases substantially. Through multiple regression analysis, predictive equations have been designed and studied in order to best predict the probability of disease in the clinically important NAFLD (i.e. NASH or advanced fibrosis).(9)

For example, the BARD score represents one of the most basic algorithms, comprising of biochemical and clinical parameters readily available to the clinician (BMI, AST/ALT ratio, presence of diabetes).(79) While the BARD score is relatively easy to use, it has been proven to be inferior to several other non-invasive biomarker panels.

A comprehensive meta-analysis by Musso et al. looked at a total of 21 non-invasive biomarker panels to identify their characteristics and diagnostic performances. Five panels were found to detect the presence of advanced fibrosis in patients with NAFLD, including BARD score, Fibrotest, enhanced liver fibrosis (ELF) panel, combined panel, and NAFLD fibrosis score. Of these five biomarker panels, Fibrotest, ELF panel, combined panel, and NAFLD fibrosis score all demonstrated increased diagnostic accuracy when compared with BARD score with test accuracy (AUROC) ranging from 0.80 to 0.90.(9) When comparing the remaining four biomarker panels for the detection of advanced fibrosis, only the NAFLD fibrosis score has been most extensively validated. Fibrotest, ELF, and combined panel are limited by lack of external validation besides the original study so their reproducibility in different population remains unknown.(9)

The NAFLD fibrosis score consists of seven routinely measured clinical and biochemical parameters including age of the patient, BMI, presence of diabetes, AST, ALT< platelets, and albumin. With these parameters on hand, the clinician can then predict either high or low probability that the patient has advanced fibrosis. A meta-analysis of 13 studies and 3604 participants demonstrated that the NAFLD fibrosis score has an AUROC of 0.85 (0.81-0.90).(9) The major limitation of the NAFLD fibrosis score, however, is that a large percentage of patients fall between the cutoffs for low or high probability of advanced fibrosis, and therefore are indeterminate for fibrosis. The same meta-analysis demonstrated that 20% to 58% of patients have indeterminate results using the NAFLD fibrosis score.(9) These patients would therefore require alternative methods of fibrosis detection.

Unlike for the detection of advanced fibrosis, biomarker panels for the detection of NASH is comparatively lacking at present. For the detection of NASH, six major biomarker panels have been studied. They include NASH Test, NASH Predictive Index, Obesity-related NASH Diagnostics, NASH Clinical Score, NAFIC score, and Plasma ELISA-detected cytokeratin-18 (CK-18).(9) Of these six biomarker panels, only the cytokeratin-18 has been externally validated from their original studies. In fact, cytokeratin-18 was validated in nine independent studies comprising of 856 NAFLD patients. Furthermore, the largest study on cytokeratin-18 fragments by the NASH Clinical Research Network (CRN) demonstrated that other routinely available parameters did not significantly improve its diagnostic accuracy.(80) Thus, the Plasma ELISA-detected cytokeratin-18 test comprises of detecting only one marker, the cytokeratin-18 fragment, which is released in the setting of hepatocyte apoptosis. A meta-analysis of cytokeratin-18 has found that the AUROC for this exam to be 0.82 for the detection of NASH in patients with NAFLD.(9)

While non-invasive biomarker panels comprise an exciting component of upcoming non-invasive assessment for NAFLD, they are limited by one important limitation. Thus far, all biomarker panels have been validated in cross-sectional studies. As a result, little is known of their diagnostic performance in monitoring disease progression and treatment response.(9) **Table 2.1** summarizes the diagnostic performances for the most well-studied serum biomarkers for liver fibrosis.

Table 2.1: Diagnostic performances of serum biomarkers for liver fibrosis

Study	Number of studies	Diagnostic endpoint	AUC (95%CI)	Sensitivity (%)	Specificity (%)
Fibrotest	2	Fibrosis ≥ F2	0.78 (0.72-0.85)	76 (70-84)	74 (69-81)
NAFLD fibrosis score	13	Fibrosis ≥ F3	0.85 (0.81-0.90)	90 (86-95)	60 (56-65)
BARD score	6	Fibrosis ≥ F3	0.78 (0.72-0.84)	72 (60-84)	64 (56-72)
Enhanced Liver Fibrosis (ELF) panel	2	Fibrosis ≥ F3	0.90 (0.84-0.96)	86 (80-91)	93 (90-96)
Fibrometer	1	Fibrosis ≥ 2	0.94 (N/A)	79 (N/A)	96 (N/A)

Adapted from: Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. Annals of medicine. 2011;43(8):617-49.

2.4.2 Imaging-based methods

Imaging-based methods make up the other large part of current non-invasive examinations for NAFLD. The preeminent and most studied imaging modalities will be highlighted in this section. **Table 2.2** summarizes the main advantages and disadvantages of the imaging-based methods discussed below.

2.4.2.1 Conventional ultrasound

Conventional ultrasound represents the most common and readily available modality in current institutions. It has been largely recognized in the qualitative assessment of hepatic steatosis. A fatty liver appears hyperechoic or brighter than surrounding structures due to the increased scatter and attenuation of fat-filled vesicles.(81) In one of the largest meta-analyses looking at the diagnostic performance of detecting simple steatosis using conventional ultrasound when compared to liver biopsy, the mean sensitivity ranged from 73 to 91%.(82) Differences in sensitivity depended largely on the presence of mild or severe steatosis, accounting for the fact that severe steatosis is significantly easier to detect than mild steatosis.

While conventional ultrasound offers the advantage of being common, readily available, and cheap, it also presents significant limitations. Detection of steatosis becomes increasingly difficult in the presence of co-existing fibrosis or inflammation. It is unable to differentiate

between simple steatosis, NASH or fibrosis.(83) Furthermore, conventional ultrasound also suffers from significant inter- and intra-observer reliability differences.(84)

2.4.2.2 Transient elastography

Transient elastography (TE) is an ultrasound-based vibration controlled technique that is currently the most validated and commonly used elastography technique globally.(85, 86) This technique involves the generation and velocity measurement of a low-amplitude shear wave within a region of interest in the liver. The measured wave velocity is then converted into measurements of liver stiffness.(87) TE has proven to be excellent for diagnosing advanced fibrosis (stages 3 and 4). A previous meta-analysis demonstrated a pooled AUROC of 0.94 with a sensitivity of 94% and specificity of 95% for detecting advanced (stage 3 and over) fibrosis.(9) Another more recent meta-analysis found that TE was good at diagnosing stage 3 fibrosis (sensitivity of 85%, specificity of 82%) and excellent at diagnosing stage 4 fibrosis (sensitivity of 92%, specificity of 92%).(88) However, concerns regarding the diagnostic accuracy of TE are raised when detecting lower stage fibrosis. In the same meta-analysis, it was found that TE only has moderate accuracy for diagnosing stage 2 fibrosis (sensitivity of 79%, specificity of 75%).(88)

Major obstacles to the successful implementation of transient elastography clinically are its rate of failure and associated unreliable results. These are most commonly as a result of obesity and operator inexperience.(89) Other cited reasons for unreliable results include recent food ingestion, ascites and heart failure. A recent study by Cassinotto demonstrated failure rate of 14.4%. Furthermore, it was found that 8.9% of cases had unreliable results.(90) To address these limitations, newer XL probes have been developed to be used in obese patients. These probes are able to assess deeper regions of interest by emitting lower central US frequencies, thus better overcoming the challenges posed by excess subcutaneous fat. Initial study of the XL probe proves optimistic, although somewhat muted results when compared to a non-obese cohort. Wong et al. demonstrated sensitivity and specificity of 78% for the detection of stage 3 or higher fibrosis in obese patients using the XL probe. The underlying caveat is that cutoff rates are different for the XL probe versus the regular probe.(91)

2.4.2.3 Acoustic radiation force impulse imaging

Acoustic radiation force impulse imaging (ARFI) comprises an alternative method of assessing liver fibrosis. In ARFI imaging, small short-duration acoustic impulses are generated in regions of interest, which cause mechanical excitation and shear wave propagation. Based on shear wave propagation velocity as picked up by the US machine, information of liver stiffness can be inferred. Similar to transient elastography, ARFI imaging can be readily available in a clinical setting with immediately available test results to facilitate workflow.(92) Furthermore, an advantage of the ARFI technology is that it can be implemented in a conventional US machine without the need for a separate Fibroscan. Therefore, to avoid multiple different studies, patients who need simultaneous conventional ultrasound evaluation for ascites or hepatocellular carcinoma screening would be more easily assessed with ARFI than transient elastography.(92)

In terms of diagnostic accuracy, the most recent study by Cassinotto et al. found that ARFI imaging performed on-par with transient elastography for the detection of stage 3 and stage 4 fibrosis.(90) Similar results were demonstrated in previous studies.(88, 93) Furthermore, failure rates were found to be lower for ARFI imaging than transient elastography. Cassinotto et al. demonstrated failures rates of only 0.7%.(90) That being said, ARFI imaging is similarly limited by unreliable results in obese patients as in transient elastography. Compared to studies using transient elastography and XL probes, studies using ARFI imaging have similar or higher unreliable rates.(94, 95) In summary, ARFI imaging proves an alternative to transient elastography. That being said, the literature on ARFI imaging is more scant than on transient elastography and more investigation in NAFLD patients is needed.

2.6.2.4 Magnetic resonance elastography

Magnetic resonance elastography (MRE) is a comprehensive method of assessing for liver fibrosis in patients with NAFLD. It uses a modified phase-contrast technique for imaging the propagation of shear waves in the liver. In general, MRE comprises an excellent ability in detecting significant (stage 3 and higher) fibrosis.(92) A recent systematic review of nine studies calculated a mean AUROC for detecting ≥ stage 3 of 0.90 and for detecting stage 4 of 0.91.(96) In another prospective evaluation, MRE was found to have an AUROC of 0.924 for discriminating advanced fibrosis (stage 3 and 4) from milder fibrosis (stage 0 to 2).(97) Finally in a separate meta-analysis, MRE was found to have a sensitivity of 92% and a specificity of 96% of distinguishing advanced fibrosis (stage 3 and 4) from milder fibrosis (stage 1-2).(98) These studies chose this specific discrimination (advanced vs. mild fibrosis) due to clinical relevancy. Patients with advanced fibrosis are specifically the ones that have the greatest risk of disease progression.

There is appeal of MRE over such imaging methods as transient elastography and ARFI imaging. At face value, it is less operator-dependent and more elegant of a modality. Furthermore, recent comparison of MRE and transient elastography for the staging of liver fibrosis demonstrated that MRE was significantly more accurate for detecting liver fibrosis stage ≥ 2 and stage 4.(99) Despite these optimistic results, further large-scale prospective trials are recommended to compare MRE from US-based elastography in NAFLD.(92) Current studies are limited by small populations. When it comes to MRE, there are also limitations in terms of costs, increased time-consumption and lack of broad availability.(9) These limitations makes MRE less appealing for routine screening purposes for NAFLD patients in clinical practice. On the other hand, its high sensitivity and specificity opens the doors to the potential of using MRE as a confirmation tool for when ultrasound-based elastography fails or when more detailed imaging is necessary.(92)

Table 2.2: Summary of noninvasive imaging modalities for NAFLD

Imaging	Advantages	Disadvantages
US	• Ready availability • Low cost	• Only qualitative assessment of steatosis

	• Provides evaluation of liver architecture	• Limited by inter- and intraobserver variability
TE	• Short processing time (<10 minutes) • Ambulatory clinical setting • Immediate results	• Limited reliability in obese individuals • False positives (ascites, congestion)
ARFI	• Readily available • Immediate results • Failure rates less than TE • Allows for simultaneous sonographic imaging of the liver	• Failed or unreliable measurements • Does not allow for quantification or assessment of steatosis
MRE	• Can be accomplished in ~20 minutes • No additional hardware needed • No contrast • Not affected by obesity • Simulatenous MRI for liver architecture and carcinoma screening	• Requires MRI facility • Results not specific to NAFLD patients • Cannot distinguish between inflammation and fibrosis • Cannot be used in some patients with implantable devices

Adapted from: Hannah WN, Jr., Harrison SA. Noninvasive imaging methods to determine severity of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Hepatology. 2016;64(6):2234-43.

2.4.3 Liver biopsy

Liver biopsy represents the gold standard in the diagnosis of NAFLD. The unmatched sensitivity and specificity of histological assessment makes it a cornerstone for the evaluation of chronic liver disease. However, by definition, liver biopsy is an invasive procedure, which carries with it real consequences. As a result, both physicians and patients may find it difficult to carry out a biopsy, especially in light of advancing alternative noninvasive methods as outlined in the previous section. (100)

From a historical point of view, liver biopsy was used almost entirely as a diagnostic tool. However, with better understanding of natural history of liver diseases as well as new therapies for patients, nowadays, it has become important in clinical management as well. The three recognized roles for liver biopsy include 1) diagnosis, 2) disease staging for prognosis, and 3) in assisting therapeutic decision making. (100) **Table 2.3** summarizes the current-day indications for liver biopsy.

Table 2.3: Indications for liver biopsy

Diagnosis
• Multiple parenchymal liver disease
• Abnormal liver tests of unknown etiology
• Fever of unknown origin
• Focal or diffuse abnormalities on imaging studies
Prognosis- Staging of known parenchymal liver disease
Management- Developing treatment plans based on histologic analysis

Adapted from: Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver D. Liver biopsy. *Hepatology*. 2009;49(3):1017-44.

Despite its universal acceptance and role in current clinical management of chronic liver disease, liver biopsy is not without its faults. Liver biopsy is associated with estimated major complication risks in 1-3% of cases, including the possibility of death in 0.01%. (101) In a separate Canadian study looking at 4275 biopsies over 10 years, liver biopsy was found to have an overall mortality rate of 0.14%. Other complications included pain requiring admission in 0.51% and bleeding in 0.35%. These complications were associated with a median cost of $ 4,579 (CAD). (102) For these reasons, liver biopsy is known to cause significant patient anxiety. (100)

Depending on the technique for procuring the biopsy specimen, there are errors related to sample location and sample size. The average liver biopsy specimen will typically only yield 0.05 cm^3 of an organ that ranges from 800 to 1000 cm^3. This corresponds to less than 1:50,000 ratio of the total volume. (101) As a result, liver biopsy has been shown to incorrectly exclude NASH in up to ¼ of cases and misclassify fibrosis stage in 1/3 of cases. (103)

In summary, liver biopsy remains the reference standard for the diagnosis of liver disease. Given the underlying risks, discussions of benefit and risk should be undertaken between physician and patient before this invasive procedure. With the advance of multiple noninvasive modalities, liver biopsy may not be the most effective diagnostic measure, for example, in the context of screening for NAFLD.

2.5 Management of NAFLD

Currently, there is widespread agreement that all patients with NAFLD should undergo lifestyle intervention in order to control metabolic risk factors such as central obesity and high fasting blood sugar. These changes include the promotion of weight loss, dietary plans, and increased physical exercise. All guidelines promote lifestyle changes as the first-line treatment for all NAFLD patients. (78)

2.5.1 Lifestyle changes

Given the significant association between central obesity and NAFLD, weight loss comprises a major part of the lifestyle intervention for patients. The exact amount of weight loss differs according to guidelines, due to a paucity of specific data related to weight loss in NAFLD. The European Association for the Study of the Liver recommends a weight loss of 7% in overweight and mildly obese patients. The American societies recommend at least 3-5% weight loss to improve steatosis and up to 10% to improve inflammation. (78)

The American recommendations were primarily based on a randomized controlled trial of 31 obese patients with NASH that looked at liver histology improvements from intensive

lifestyle changes versus basic education alone. (104) Intensive lifestyle changes were defined as a plan consisting of diet, behavior modification and 200 minutes/week of moderate physical activity for 48 weeks. The authors found that intensive lifestyle changes led to a weight loss of 9.3% versus 0.2% in the basic education arm. Furthermore, there was significant improvement in steatosis, necrosis and inflammation on post-treatment biopsies. Importantly, no improvement in liver fibrosis was found in the either treatment arm. (104)

In terms of alimentation, all guidelines currently recommend avoidance of heavy alcohol consumption in NAFLD patients. (78) Similarly, all societies recommend a hypocaloric diet for the promotion of weight loss. The diet should be low in carbohydrate and saturated fats with specific avoidance of fructose-enriched drinks. Diets rich in fiber, anti-oxidant rich fruits and vegetables are recommended. (78)

Finally all guidelines strongly recommend implementation of a physical exercise regime. (78) Such exercise regimes should comprise of at least 150 min per week of moderate-intensity physical activity. These regimes are based on MR spectroscopy studies looking at the effect of exercise alone without diet modification. In these studies, exercise regimens typically involved 2-3 sessions per week of 30-60 minutes physical activity over 6 to 12 weeks. These studies demonstrated significantly decreased liver fat content without change to overall weight. (105-107)

2.5.2 Pharmacologic therapy

Currently, pharmacologic therapy is recommended to be used only in cases of biopsy-proven NASH. (78) There are two large classes of medications that have been tested in the treatment of NAFLD. They include insulin sensitizing agents such as metformin and thiazolidinediones, and antioxidants such as vitamin E. Other miscellaneous agents such as ursodeoxycholic acid (UDCA) and Omega-3 fatty acids have also been studied.

2.5.2.1 Insulin-sensitizing agents

Metformin was one of the first agents studied in the treatment of patients with NASH. The earliest studies looking at the effect of metformin on liver enzymes and liver histology in NASH patients demonstrated decrease in insulin resistance and liver aminotransferases. However, they did not demonstrate any significant improvement in liver histology, namely inflammation. (108, 109) Later studies also supported this finding with no significant improvement of metformin on liver histology. (110) Furthermore, other studies failed to demonstrate major improvements on insulin resistance or aminotransferase levels. (111) A definitive randomized control trial of metformin versus placebo by Haukeland et al. concluded no significant difference between the two branches. (112) Given these findings, metformin is not recommended as a treatment for patients with NASH. (5)

In contrast to metformin, thiazolidinediones have demonstrated more positive results in the treatment of NASH. Thiazolidinediones, which consist of pioglitazone and rosiglitazone, are a specific class of drugs that promotes adipogenesis and fatty acid uptake peripherally. While randomized controlled trials involving rosiglitazone have been met with mixed results, pioglitazone has proven to be effective in the treatment of steatohepatitis. (5) In a randomized controlled trial of patients with NASH given 45 mg/day of pioglitazone versus placebo, it was demonstrated that there was significantly improved aminotransferases, histologic steatosis, ballooning and inflammation. There was improvement of inflammation in 73% of patients treated with pioglitazone versus 24% in the placebo-treated arm. (113) Another randomized controlled trial demonstrated significant improvement of hepatocellular injury and fibrosis. (114) The large PIVENS study also found that significantly higher number of patients treated with pioglitazone demonstrated resolution of NASH versus those treated with placebo. (115) These findings were repeated in a recent meta-analysis of five randomized controlled trials, which concluded significant improvement on steatosis and inflammation with pioglitazone. (33) Given the literature, pioglitazone is currently recommended in the treatment of biopsy-proven NASH by both American and European guidelines. (78)

There are a few potential drawbacks associated with the use of pioglitazone. Firstly, it has no proven effect on histologic liver fibrosis. (33) Secondly, there is currently controversy surrounding possible long-term safety effects associated with thiazolidinedione use. These safety hazards include increased risk of cardiovascular disease, congestive heart failure,

bladder cancer, and osteopenia. (5) A meta-analysis of 16,390 patients with type 2 diabetes on pioglitazone treatment found that there was a slightly higher rate of CHF with pioglitazone use (2.3% versus 1.8% in the control arm). At the same time, however, it was also found that overall mortality and rate of myocardial infarction and stroke was significantly reduced. (116)

2.5.2.2 Antioxidant agents

Given that oxidative stress is considered a key factor in hepatocellular injury and pathogenesis of NASH, antioxidants comprise the other large category of NASH treatment options. In particular, Vitamin E has been most studied. (115, 117) Similarly to pioglitazone use, vitamin E has been associated with decreased aminotransferase levels, improvement in liver histology (namely, steatosis, inflammation and ballooning), with no significant effect on liver fibrosis. (5) The previously mentioned PIVENS study also found that vitamin E was associated with significantly higher rate of improvement in NASH when compared with the placebo arm. (115) Given these findings, vitamin E comprises the other medication recommended for the treatment of biopsy-proven NASH. (78)

As with pioglitazone, there is currently some uncertainty surrounding the regular use of vitamin E. In particular, high-dose vitamin E has been associated with increased all-cause mortality. It has also been associated with increased hemorrhagic stroke and prostate cancer risk. (118, 119) That being said, further research is needed as other studies have failed to confirm the association between vitamin E and increased mortality. (120)

2.5.2.3 Miscellaneous agents

Other agents such as ursodeoxycholic acid and omega-3 fatty acids have been studied in the treatment of NASH. Thus far, ursodeoxycholic acid has not shown histologic improvement over placebo in randomized control trials. (121) Omega-3 fatty acids, used in the treatment of hypertriglyceridemia, is currently being investigated in the treatment of NASH with an ongoing multicenter trial ongoing in the United States. (5) New

pharmacologic treatment in NASH is an evolving field with further research needed to fill this knowledge gap.

2.6 Areas of uncertainty

There are several areas of uncertainty in the study of NAFLD. The first area of uncertainty is in the natural history of NAFLD. The bulk of the current literature on natural history is limited by relatively short-term observational studies with composite outcomes that do not differentiate well between hepatic, metabolic, and cardiovascular complications.(4) Given the chronicity of NAFLD, there is still the need for definite, well-controlled, longitudinal studies over the long-term.(51)

Another area of uncertainty revolves around current noninvasive modalities for the diagnosis of NAFLD. Current generation examinations such as transient elastography and magnetic resonance elastography show promise for the diagnosis of advanced fibrosis.(9) Biomakers such as cytokeratin 18 levels (CK-18) have demonstrated good correlation with the presence of NASH.(122) Nonetheless, these examinations have not entered into routine clinical practice, partly because of the need for performance assessment in different populations, longitudinal evaluations over longer term, as well as efficacy in the setting of treatment response.(123)

Similarly in the treatment of NAFLD, there still exist pertinent gaps of knowledge. The current management of NAFLD is a multidisciplinary approach with lifestyle change and weight loss supplemented by pharmacologic therapy in cases of NASH.(124, 125) Nonetheless, there is a real need for more long-term multicenter randomized controlled trials, in particular regarding the long-term efficacy and adverse effects associated with current pharmacologic therapies.(124, 125) The development of a highly effective and specific treatment for NAFLD is needed.(123)

Finally, as stipulated by multiple liver societies around the world, there has yet to be any study looking at the cost-effectiveness of screening for NAFLD in the general population or high-risk groups.(78)

2.7 Current recommendations for screening

At present, there is much debate among the different international organizations relating to the decision to screen for NAFLD in either the general population or high-risk populations affected by diabetes or obesity. The European Association for the Study of the Liver (EASL) has brought out a statement in 2009 in support for screening in high-risk groups.

> *Screening for NAFLD/NASH is not recommended in the general population; it is recommended in patients with metabolic risk factors and/or well characterized insulin resistance.(1)*

The EASL underlines the need for non-invasive quantification of fibrosis and steatohepatitis in order to allow screening of large numbers of at-risk patients without the need for biopsy. They recommend correlation of elastometry with serum markers for fibrosis and also outline the need for accurate non-invasive diagnosis of steatohepatitis.(1)

Similar beliefs are underscored with other liver disease organizations in the Asian Pacific and China.(21, 126) According to the Asia-Pacific Working Party on NAFLD:

> *[…] patients with suspected NAFLD should undergo baseline tests that allow definition of NAFLD (discussed earlier in relation to proposal 1), identification of the underlying metabolic factors, exclusion of other disorders, and assessment of the likely severity of NAFLD/NASH. These tests encompass biochemical and hematological indices, anthropometry, hepatic imaging, and determination of insulin sensitivity.(126)*

Contrary to these organizations, however, the American Gastroenterological Association (AGA), American Association for the Study of Liver Diseases (AASLD), and American College of Gastroenterology (ACG) do not recommend screening at this time.

> *Screening for NAFLD in adults attending primary care clinics or high-risk groups attending diabetes or obesity clinics is not advised at this time due*

> *to uncertainties surrounding diagnostic tests and treatment options, along with lack of knowledge related to the long-term benefits and cost-effectiveness of screening.*(5)

The American associations argue that there are still significant gaps in knowledge relating to the natural history and treatment of NAFLD. Simple biochemical tests such as AST and ALT are unreliable in patients with NAFLD and NASH, as they are oftentimes normal even when there is disease. On the other hand, more sensitive and advanced tests such as ultrasound elastography are considered more expensive and time-consuming. That being said, they believe that more research is needed regarding the long-term costs and benefits of screening.(5)

Chapter 3

Criteria for screening

3.1 Introduction to screening

The central idea of early disease detection and treatment is essentially simple. However, the path to its successful achievement (on the one hand, bringing to treatment those with previously undetected disease, and, on the other, avoiding harm to those persons not in need of treatment) is far from simple though sometimes it may appear deceptively easy.(6)

The main objective of screening is to uncover those suffering from disease among a population of apparently healthy individuals. Upon discovery, the diseased can then be treated. Advantages to screening is firstly, the treatment of otherwise undiscovered populations. Secondly, if the disease is communicable, screening may help prevent further spread of said illness. Early detection of illness is therefore beneficial not only to individuals but society as a whole. Over the long term, society would benefit from a more healthy and productive population. (6)

Conceptually, screening is a byproduct of a developed healthcare system. In developing regions of the world where there is a large burden of overt and communicable disease, most of the resources are dedicated to treating recognizable disease. There are few resources available to the allocation of screening programs. On the contrary in developed nations, where communicable diseases such as infectious diseases are less important, more insidious chronic diseases have become the forefront of the healthcare burden. Furthermore, by definition in developed countries, there is ample resource available to the allocation of screening programs. Thus, it is recommended that the practice of screening for disease in developed countries should be paramount and widespread. That being said, not all chronic illnesses can be screened, nor does it make sense to screen all chronic illnesses. Specific criteria should be met before the initiation of such endeavors. (6)

3.2 Basic definitions

According to the Commission on Chronic Illness Conference on Preventive Aspects of Chronic Disease in 1951, screening is officially defined as “the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedure which can be applied rapidly. Screening tests sort out apparently well persons who probably

have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment". (127)

A few important points should be highlighted in the above statement. Firstly, the detection of insidious disease should be applied rapidly. A long and laborsome screening test is unlikely to work well. Secondly, unlike diagnosis, screening can function on probabilities. While a screening test may be diagnostic or specific, it more importantly has to be a sensitive test. Once suspicious findings are found on screening, further confirmatory diagnostics can be undertaken.

Screening can be either large-scale or more selective. Large-scale screening or *mass screening* is screening on the largest scale, where no pre-selection process is undertaken first. On the contrary, *selective screening* refers to screening in high-risk groups in the population. This latter is usually more cost-effective. As a relevant example, screening for NAFLD would intuitively make more sense in high risk diabetes and obese patients. (6)

Screening may involve a single test or a combination of multiple screening tests. Wilson and Jungner define *multiple or multiphasic screening* as "the application of two or more screening tests in combination to large groups of people". (127)

Finally, Wilson and Jungner also define the often confused term *surveillance*. In their report, they use surveillance to convey a long-term process of screening, while screening can be thought of as cross-sectional and short-term operations. In reality, however, screening and surveillance are used often interchangeably. (6)

3.3 Criteria for appraising screening

Wilson and Jungner contributed to a set of criteria to assess the validity and viability of screening. In order for NAFLD screening to be successful, these criteria need to be satisfied. **Table 3.1** summarizes the ten criteria set by Wilson and Jungner on appraising the validity of a screening program in the context of NAFLD.

Table 3.1: Applying Wilson and Jungner's criteria to NAFLD

Criteria for appraising validity of a screening program	NAFLD	NASH
1. The condition being screened should be an important health problem	✓	✓
2. The natural history of the condition should be well understood	✓	✓
3. There should be a detectable early stage	✓	+/-
4. Treatment at an early stage should be of more benefit than at a later stage	✓	✓
5. A suitable test should be devised for the early stage	✓	+/-
6. The test should be acceptable	✓	✓
7. Intervals for repeating the test should be determined	?	?
8. Adequate health serve provision should be made for extra clinical workload resulting from screening	?	?
9. The risks, both physical and physiological, should be less than the benefits	✓	✓
10. The costs should be balanced against the benefits	?	?

Additional columns for NAFLD and NASH were added for summary and illustration purposes. Adapted from: Wilson JMG, Jungner G. Principles and practices of screening for disease. Geneva, Switzerland: *World Health Organization*; 1968. Report No.: Public Health Papers No. 34. Available from: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf.

3.3.1 Importance of screened condition

In order for screening to make fundamental sense, the target disease should have serious consequences on society.(128) This way, there is easy justification from a cost-effectiveness and emotional standpoint for screening.

NAFLD is now recognized as the most common liver disease in developed nations. With the rise of obesity and diabetes, this number is only expected to augment.(2) As NAFLD progresses towards advanced liver disease, cirrhosis and hepatocellular carcinoma, the underlying costs on society will also increase.(129) Thus, this first criteria is easily satisfied in NAFLD, which has serious implications on Western society.

3.3.2 Recognizable latent symptomatic stage

A screened disease should have a detectable preclinical or latent symptomatic phase in order to justify the costs of screening.(128) The natural history of a disease can be divided into preclinical and clinical phases. Usually, the preclinical phase ends when the patient begins to have symptoms and pursues medical care. These factors depend on the population's education of the disease and of the patient's access to healthcare.(130) Thus, the detectable preclinical phase is the interval of time when the disease is able to be detected by screening. **Figure 3.1** illustrates the natural history of disease and the important phases involved in the detection and treatment of disease.

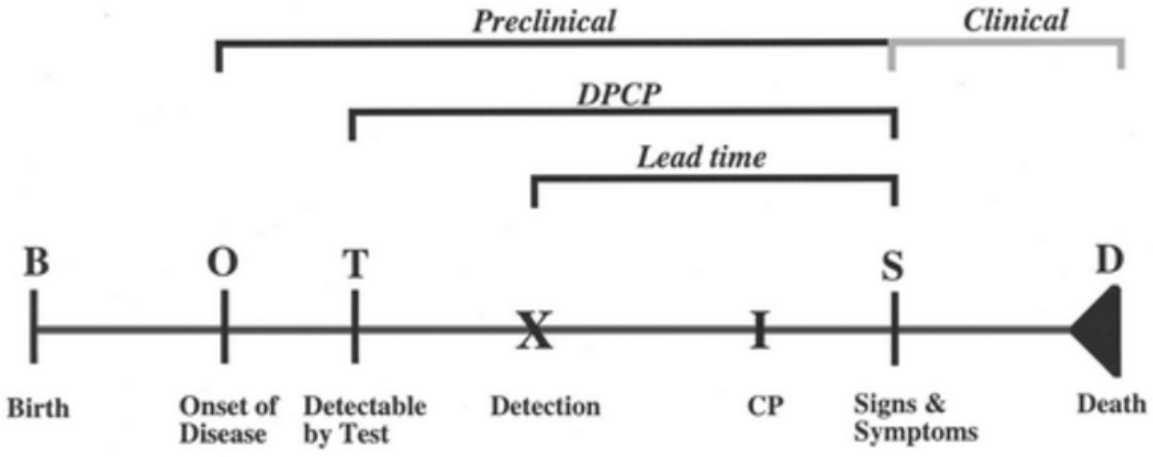


Figure 3.1: Natural history of disease and detection. The preclinical phase of disease begins with undetectable onset of pathogenesis and ends at the appearance of signs and symptoms. The clinical phase is characterized by the manifestation of signs and symptoms. The detectable preclinical phase (DPCP) is the period of time during the preclinical phase where the disease is able to be detected by a test. The lead time is defined as the interval of time ranging from the moment of disease detection and the appearance of signs and symptoms. In order for screening to be effective, the critical point needs to occur during the DPCP. Obuchowski NA, Graham RJ, Baker ME, Powell KA. Ten criteria for effective screening: their application to multislice CT

screening for pulmonary and colorectal cancers. *AJR American journal of roentgenology*. 2001;176(6):1357-62.(131)

NAFLD fits this criteria well. It represents a spectrum of disease ranging from steatosis, to steatohepatitis, to fibrosis, and finally end stage liver disease. In NAFLD, the latent or preclinical phase would be the steatotic, inflammatory and early fibrotic phase of disease, given that many of such patients are asymptomatic.(17, 51, 68, 132) Therefore, these stages of latent disease serve as ideal targets for screening.

3.3.3 Well-understood natural history

The underlying natural history of a disease should be sufficiently understood prior to any screening implementation. In particular, a critical point of the natural history of a disease should be identified. The critical point of a disease is defined at the time point whereby treatment is most effective before this point and least effective after this point. If the critical point lies before the detectable preclinical phase, then screening would be ineffective. Also, if the critical point appears shortly after the detectable preclinical phase begins, then there would be too little time to treat the patient effectively and screening would be too late. Lastly, if the critical point occurs during the clinical phase after the onset of symptoms, then there would be no sense in a screening program. (130)

While there is much to understand about the natural history of NAFLD, sufficient evidence points to the stage of inflammation or NASH as the critical point of disease. Patients with biopsy-proven NASH are significantly more likely to progress onto fibrosis and end-stage liver disease, as well as die from liver-related causes.(56, 133, 134) In the simple steatosis phase, there is still ample time to make lifestyle adjustments to prevent the progression of disease. Similarly in patients with NASH, there is evidence to suggest that both lifestyle modifications and pharmacologic agents are capable of slowing or deterring the progression of disease.(5, 78) If a screening program can target the steatotic and inflammatory phases of NAFLD, then it can be effective. Therefore, this screening criteria is satisfied in NAFLD.

3.3.4 Suitable screening test

The screening test ideally should have high diagnostic accuracy for detecting the disease in the preclinical phase. In order to detect more true-positives than false-positives when the prevalence of disease is ≤ 5%, a screening test must have a sensitivity >95% if the specificity ≤95%. Given that most screening test do not have such exigent diagnostic performances, screening programs therefore absorb the costs of the false-positive results.(130) However, as specified by Wilson and Jungner, screening by definition does not need to be diagnostic, but rather guide the clinician to further steps in diagnosis.

Furthermore, given the large target populations for screening, suitable screening tests should ideally be non-invasive.(6) It is important to keep in mind that at the time of screening, individuals are at minimal risk of morbidity or morality. Given that large number of patients are screened, even the smallest adverse effect can likely offset any substantial benefit.(130)

In NAFLD, there are currently an array of noninvasive biomarker panels and imaging-based modalities available for the detection of steatosis, steatohepatitis, and fibrosis. While their diagnostic performance may vary and are still inferior to the reference standard of liver biopsy, they boast the advantage of rapidity, accessibility, and non-invasiveness.(9, 92, 135) Furthermore, recent algorithms have been developed to combine these different noninvasive tests into multiphasic screening algorithms to improve diagnostic accuracy.(9)

3.3.5 Acceptable treatment for disease

Any screening program for a disease requires a proven and effective treatment for the disease that will improve patient outcomes. This is because screening alone cannot be cost-effective. For example, many important diseases without effective treatment such as Parkinson's disease or Alzheimer's disease cannot currently be screened cost-effectively.(130)

The treatment should ideally not have too many adverse effects, which would offset any long-term benefits. Given the early onset of treatment in a screened population, long-term toxic effects should be well-studied. Since most people place greater value on the next few

years of their life rather than in the far future, any negative effect from treatment would more profoundly diminish a patient's quality of life.(128)

In NAFLD, lifestyle modifications including weight loss, diet change and exercise are proven and benign treatment towards halting the progression of disease. In terms of pharmacologic treatment, pioglitazone and vitamin E are at the forefront of suitable agents currently recommended for the treatment of biopsy-proven NASH. However, more research is still needed in terms of pharmacologic agents for the treatment of NAFLD. Both pioglitazone and vitamin E are mired in controversy regarding potential long-term side effects.(5, 78) Nonetheless, the fact that lifestyle modifications remain the primary treatment option in early stages of NAFLD means that this screening criteria is satisfied.

3.3.6 Cost-effectiveness of screening

According to Wilson and Jungner, the cost of case-finding (including the cost of diagnosis and treatment) should be balanced in an economic sense with total healthcare expenditure. Governments should recognize that screening is a continuous process, spanning many years rather than just a one-and-done project. As such, prior to initiating a screening program, the overall cost-effectiveness should be evaluated in terms of benefits (i.e. additional life years or quality-adjusted life years) to the patient and dollar cost to the healthcare system. As well, in order for a screening program to be successful, there requires adequate health service provisions be made. In developed countries, this means allocation of enough healthcare resources towards the extra clinical workload and strain on their healthcare systems.(6, 130)

These criteria represent a current knowledge gap in NAFLD. Thus far, there has been no cost-effectiveness study looking at the cost-effectiveness of NAFLD screening in Western countries. Issues to resolve are first, is screening (including the subsequent diagnosis and treatment of positive cases) cost-effective? Secondly, what is the optimal interval of screening? Finally, if screening is cost-effective, do governments have the adequate resources to move towards a screening program? A study evaluating the long-term cost-

utility of screening in the general population and at-risk groups demonstrating metabolic risk factors is currently needed.

3.4 Biases in screening

The lead time is defined as the time between disease detection and usual clinical presentation and diagnosis. The intention of screening is to diagnose a disease earlier than the usual clinical presentation. However, early diagnosis may not always prolong a patient's survival. This is the case when the critical point for treatment does not occur during the detectable preclinical phase. Thus, lead time bias occurs when comparisons of survival are not adjusted for the timing of diagnosis.(130) For example, if a patient is genetically predisposed to die by a certain age, it would not matter if the patient was screened as a baby and treated early or detected by usual clinical presentation and treated later. If survival was calculated by timing of diagnosis, then there would be a lead time bias between the screening and no screening. This concept is illustrated in **Figure 3.2**, which compares screening and no screening in a situation where there is no early critical point in treatment and survival.

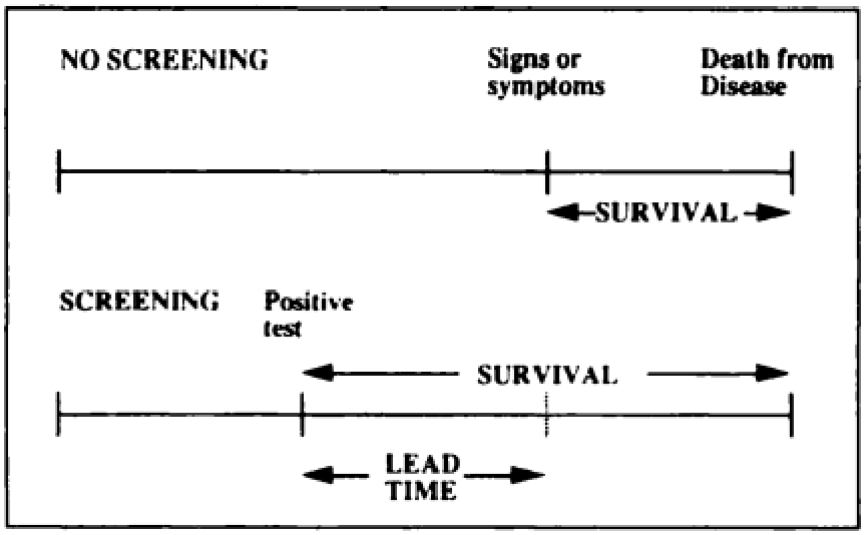


Figure 3.2: Lead time bias. Earlier diagnosis has no effect on the timing of death. There is no critical point at which treatment is more effective during the detectable preclinical phase, thus the overall survival is equal between screening and no screening. Black WC, Welch HG. Screening for disease. *AJR American journal of roentgenology*. 1997;168(1):3-11.

Length time bias is a form of selection bias that occurs when comparisons between screening and not screening does not account for differing rates of disease progression. The probability that a case will be detected by screening and the length of its detectable preclinical phase are proportional.(130) Therefore, diseases that progress slower are more likely to be detected by screening than diseases that progress faster. For example, patients with slow-growing cancer are more likely to be detected by screening than fast-growing cancer. Thus, the same patients who are detected by screening in the preclinical phase would on average do better than patients who were detected from symptoms in the clinical phase. This would give the overall impression that a screening program is effective, when in fact, only the slow-growing and less dangerous cancer is being selected in the first place. **Figure 3.3** illustrates the concept of length time bias.

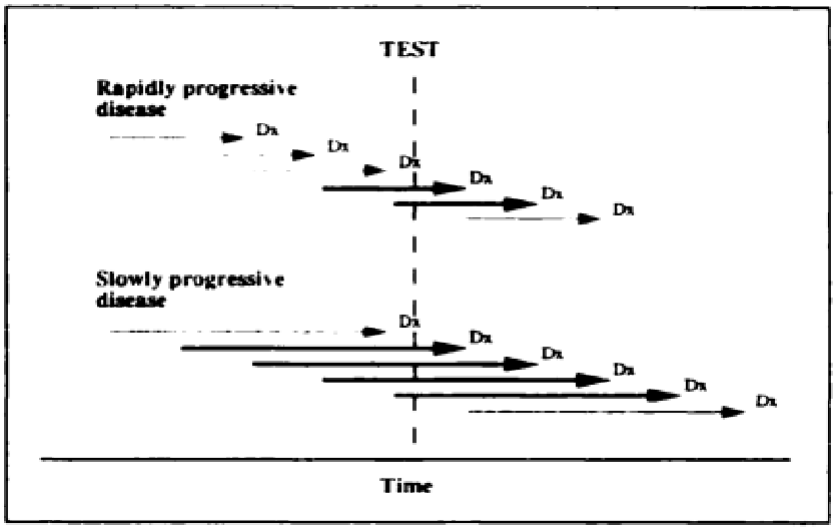


Figure 3.3: Length time bias. The overall chance of detection is directly proportional to the rate of disease progression. The length of each arrow illustrates the length of detectable preclinical phase until the usual clinical diagnosis (Dx). In this example, testing at the same point in time would detect two rapidly progressive disease cases versus four slowly progressive disease. Black WC, Welch HG. Screening for disease. *AJR American journal of roentgenology*. 1997;168(1):3-11.

The final bias involved with screening is called overdiagnosis bias. This bias occurs through the overestimation of survival duration in screen-detected cases are a result of pseudodisease inclusion. Pseudodisease refers to the subset of the subclinical disease population that does not become overt disease before they die of other causes. One of the

most well-known example would subclinical prostate cancer, which occurs in the majority of men, but ultimately only contributes to the death of a minority.(130)

Depending on the knowledge of the natural history of disease, it may be difficult and sometimes impossible to determine the individual contributions of these biases on a screening program. However, evaluators should be aware of the positive effect that lead time, length time, and overdiagnosis biases have on the overall effectiveness of a screening program. Ultimately, the apparent effect of screening should be thought of as the sum of the real effect and the positive bias related with earlier diagnosis.(130)

Chapter 4

Economic analysis in healthcare

4.1 Rationale of health economic analyses

In the developed world, the burden of chronic illness on the health care system is increasing dramatically. The rise in prevalence of heart disease, hypertension, diabetes and obesity, coupled with an aging population, is contributing to significant increases in healthcare costs.(27) **Figure 4.1** demonstrates the trend in health expenditure in Canada according to the Canadian Institute for Health Information.(136) The total health expenditure in Canada is forecasted to grow by 2.7% in 2016. Since 2010, the rate of growth in healthcare spending is being outpaced both population growth and inflation, accounting for 10% of Canada's GDP. Similar circumstances are seen in other Western countries, although only United States, France, Germany, and Sweden spend more of their GDP on healthcare expenditure.(136) In such an environment, governments need to be prudent in their budget expenditure and attempt to make informed evidenced-based decisions. Economic analyses serve as an increasing important tool in the appraisal of healthcare programs.

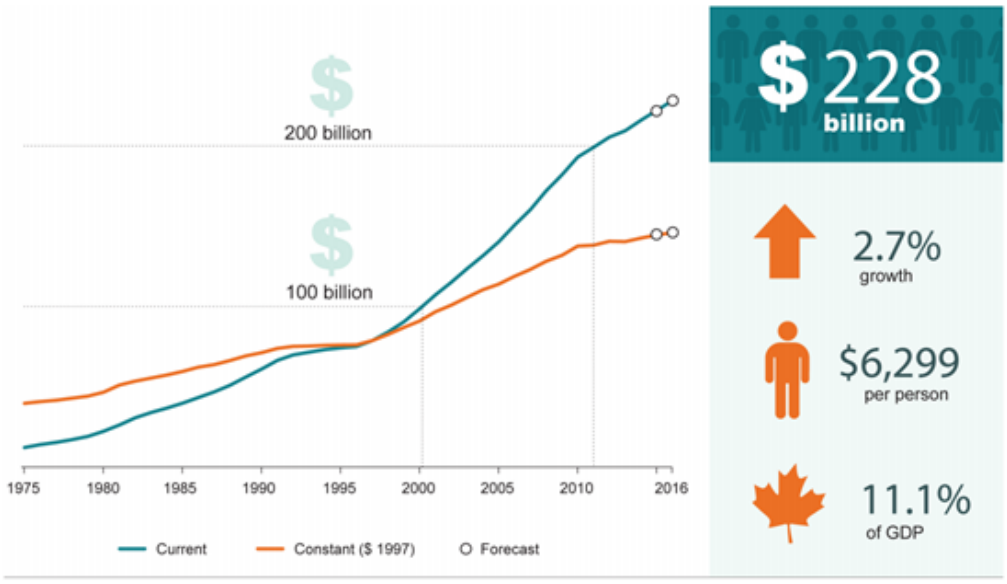


Figure 4.1: Trend in total health expenditure in Canada. Healthcare spending is projected to reach $228.1 billion in 2016. This represents a staggering 11.1% of Canada's GDP and equals $6,299 per Canadian. The total health expenditure growth in 2016 is projected to be 2.7%. *Canadian Institute for Health Information*. National Health Expenditure Trends, 1975 to 2016. (Ottawa, Ont: CIHI; 2016). Accessable from: https://www.cihi.ca/sites/default/files/document/nhex-trends-narrative-report_2016_en.pdf

4.2 Definitions

The main objective of an economic evaluation is to "identify, measure, value and compare the costs and consequences of alternative [courses] being considered".(137) In health economics, there are different types of costs and consequences. Many costs are fixed and well-defined, such as dollar costs related to medications or medical equipment. Other costs can be variable and indirect, such as for example, the cost of time lost to patients related to an intervention. Consequences or outcomes can vary according to the type of economic analysis. They can include number of life years gained, number of deaths avoided, and quality-adjusted life-years.(137)

A key concept to understand in this area of economics is that of *opportunity cost*. In a limited-resource environment such as the Canadian healthcare system, choices need to be made regarding which intervention can be provided for which situation.(137) The decision to fund one healthcare intervention may mean that another cannot be funded. Thus, the opportunity cost of funding one intervention is the health benefits of another intervention that inevitably was not funded. **Figure 4.2** demonstrates in diagrammatic form the concept of opportunity cost.

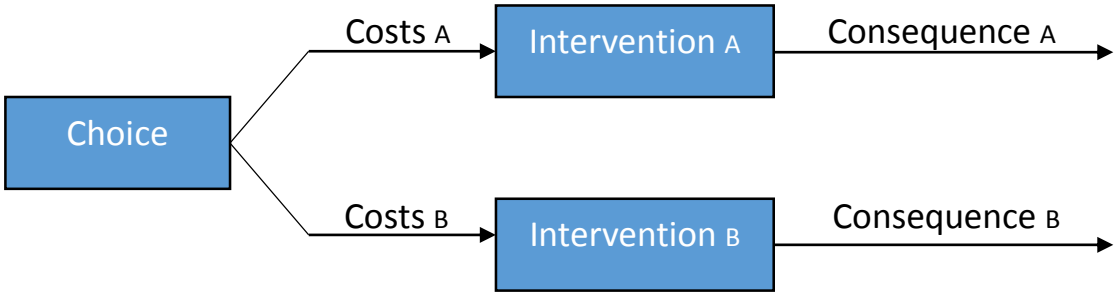


Figure 4.2: Illustration of opportunity cost. A key concept in economics, opportunity cost refers to the cost incurred by choosing between mutually exclusive alternatives. If intervention A is chosen, then the costs and benefits of intervention B is relinquished. Thus, it is important to make sure that the most cost-effective intervention is determined prior to the choice.

An advantageous characteristic of economic evaluation is that when comparing two alternative courses of action, it is not necessary to compare every possible cost and consequence under the sky.(137) Instead, the decision on the best course of action can be made by comparing the *incremental changes* between alternative interventions. Analysis of

incremental rather than total change allows for more streamlined but equally effective evaluations.

4.3 Types of studies

The five types of economic evaluations include cost-consequence analysis (CCA), cost-minimization analysis (CMA), cost-utility analysis (CUA), cost-effectiveness analysis (CEA), and cost-benefit analysis (CBA). The choice of appropriate type of evaluation depends on the research question, including the target population, the interventions/comparators, and the availability of outcome data.(137) **Table 4.1** summarizes the key differences between these types of economic evaluations.

4.3.1 Cost-consequence analysis

In cost-consequence analysis (CCA), individual costs and outcomes of comparators are listed separately in a disaggregate fashion.(137) The costs can include for example, costs of medications, costs of hospitalizations etc. The outcomes can include for example, deaths avoided, major side effects, impact on quality of life. The main advantage of this type of study is that it can obtain a general idea of the overall impact between comparators. The main disadvantage is that since the intervention outcomes are not presented in a global fashion, the reader of the study must weight and value each individual component by themselves.

4.3.2 Cost-minimization analysis

The goal of this type of study is to identify the least costly intervention between comparators that have identical outcomes. Usually, cost-minimization analyses (CMA) are undertaken when two comparators have been proven to have the same efficacy, for example through an equivalency or non-inferiority clinical trial. Thus, in these scenarios, the main outcome measures is the difference in cost between comparators, in order to select the least costly option.(137)

4.3.3 Cost-effectiveness analysis

In this type of economic evaluation, the outcome measure is in natural units such as life-years gained, lives saved, or clinical events avoided. According to the Canadian Agency for Drugs and Technologies in Health, it should be the Reference evaluation of choice when cost-utility analysis is inappropriate.(137) The main advantage of a cost-effectiveness analysis (CEA) is that it is a more straightforward analysis when compared with a cost-utility or cost-benefit analysis. However, the reliance of CEA on narrow-scoped natural health outcomes can be disadvantageous. CEA results of one intervention can only be compared with results from other interventions when the outcomes are the same. Sometimes, the outcome measure (i.e. clinical events avoided) may not account for the all of the important outcomes due to an intervention.(137)

4.3.4 Cost-utility analysis

Outcome measures in cost-utility analysis (CUA) are measured in health-related preferences, often expressed as quality-adjusted life years (QALYs) gained. Cost-utility analyses (CUA) are particularly useful when interventions impact the health-related quality of life (HRQL) and should be used as the Reference method of choice in those cases.(137) CUAs are limited, however, by the availability of health utility information. Furthermore, health utility data may defer according to the method or instrument used to collect the information.(137)

4.3.5 Cost-benefit analysis

Cost-benefit analyses (CBA) differ from other types of economic analyses in the sense that it values all costs and outcomes in monetary terms. These monetary values are obtained usually through a willingness-to-pay approach.(137) Although CBAs theoretically would best address the question of resource allocation given their monetary outcomes, their methodological difficulties limit their application in reality. In particular, there are methodological and ethical issues with relating health outcomes in monetary terms. CBAs may be appropriate, however, in certain situations where an intervention outcome is difficult

to quantify in QALYS (i.e. short-term symptom relief) or where an outcome variable is difficult to value using any health outcome (i.e. more convenient dose form).(137)

Table 4.1: Summary of different types of economic analyses

Analysis type	Cost measure[1]	Comparison type[2]	Outcome measure
Cost-effectiveness	$	÷	Natural units
Cost-consequence	$	Vs.	Natural units
Cost-utility	$	÷	Utilities (i.e. QALYs)
Cost-benefit	$	÷ or -	$
Cost-minimization	$	Vs.	Assume same

1. Any currency
2. ÷ (Ratio); Vs. (Comparison); - (Net summary)

Cost measure is in any currency. For cost-benefit analyses, the comparison type is in either cost-benefit ratio or net sum of costs and benefits. Adapted from: U.S. National Library of Medicine [Internet]. Maryland: Health Services Research & Public Health; 2014. [updated 2017 March 07; cited 2017 Feb 18]. Available from: https://www.nlm.nih.gov/nichsr/hta101/ta10107.html. (138)

4.4 Designing a health economic study

4.4.1 Clinical impasse and comparators

All economic analyses should begin with a clear definition of the comparators. Usually, there is a key clinical question or impasse between two or more interventions that needs to be answered with some degree of uncertainty. The reference comparator is usually the current best practice or standard of care. New interventions are potentially viable alternatives to the standard of care. In theory, they would replace the current best practice should they be found to be more cost-effective.(137)

4.4.2 Perspective

Before undertaking any economic evaluation, the perspective of the study should be properly defined. An economic evaluation may take the perspective of a healthcare system, a third party payer, a single hospital center, a patient, or even an entire society. Perspective is important because it defines the costs and consequences of the study. For example, a publicly funded healthcare system would be interested in direct costs paid by the system such as drugs, hospitalization services, and healthcare providers and staff. In contrast, taking the perspective of society would entail accounting for productivity costs related to work absenteeism and cost to employer for hiring replacement workers. Ultimately, the perspective chosen should reflect the study question at hand. According to the guidelines set by the Canadian Agency for Drugs and Technologies in Health, health economic evaluations should prioritize the perspective of the healthcare system.(137) The Institut national d'excellence en santé et en services sociaux (INESSS) recommends a societal perspective.(139) **Figure 4.3** illustrates the concept of perspective in economic evaluations.

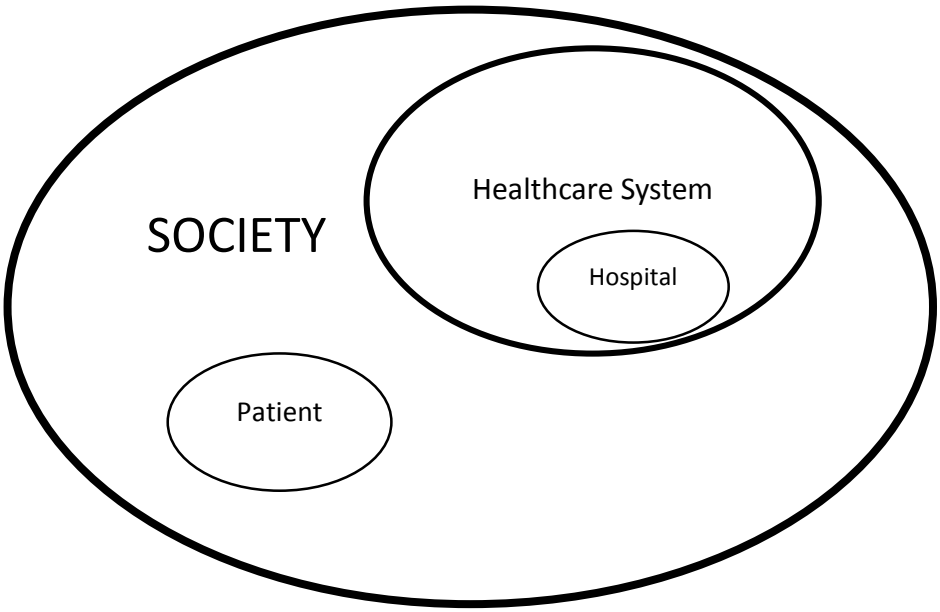


Figure 4.3: Illustrative representation of perspectives in economic evaluations. It is important to specifiy the study perspective because the latter essentially defines the basis of analysis and determines the relevant costs.

4.4.3 Costs

Costs can be divided into direct and indirect costs. In healthcare terms, direct costs represent all monetary value of goods and services consumed in the intervention, treatment, or management of patients. Direct costs can also include non-healthcare costs, such as that which is connected with care provided by family members or costs related to transportation. Indirect costs are those costs related to absenteeism, early retirement, lower productivity at work, or impaired leisure activities. Indirect costs can also be thought of as productivity losses.(137) **Table 4.2** describes the different types of costs as it pertains to perspective.

Table 4.2: Different types of costs and their associated perspectives

	Reference Case	Non-Reference Case			Examples
	Public Health Care Payer[a]	Private Payer[b]	Broader Government Payer	Societal	
Types of Costs					
Costs to publicly funded health care payer	✓		✓	✓	• Drugs, medical devices, procedures • Equipment, facilities, overhead • Health care providers • Hospital services • Diagnostic, investigational, and screening services • Informal caregivers' health care costs • Rehabilitation in a facility or at home[c] • Community-based services, such as home care, social support[c] • Long-term care in nursing homes[c]
Costs to private insurer		✓		✓	• Drugs, medical devices (falling outside of public payer) • Aids and appliances • Alternative care (e.g., chiropractic services, massage therapy, homeopathy) • Rehabilitation in a facility or at homec • Community-based services, such as home care, social support[c] • Long-term care in nursing homes[c]
Costs to government payer (beyond health care)			✓	✓	• Social services, such as home help, meals on wheels[c] • Affordable housing • Education
Costs to patients and informal caregivers				✓	• Out-of-pocket payments (e.g., copayments for drugs, dental, assistive devices) • Cost of travel, paid caregivers • Premiums paid to private insurers • Patient's time spent for travel and receiving treatment
Productivity costs				✓	• Lost productivity due to reduced working capacity, or short-term or long-term absence from work • Lost time at unpaid work (e.g., housework) by patient and family caring for the patient • Costs to employer to hire and train replacement worker
Types of Outcomes					
Health effects relevant to patients and informal caregivers	✓	✓	✓	✓	• Health-related quality of life • Life-years gained • Clinical morbidity
Non-health effects relevant to patients and informal caregivers			✓	✓	• Information available to patients • Reduction in criminal behaviour • Better educational achievements

[a] Any spillover impacts should be handled in a non-reference case analysis.
[b] Researchers should consult the private payer to determine costs and outcomes of relevance.
[c] Some of these costs may be incurred by the publicly funded health care payer, depending on the precise nature of these costs and the relevant jurisdiction.

Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada [4th Edition]. Ottawa, 2017:30.

4.4.4 Time horizon

Time horizon refers to the total time during which an economic analysis takes place. In general, it is recommended that a time horizon encompass all significant differences in costs and outcomes between comparators. Once there are no meaningful difference between costs and outcomes, there is no further point in extending time horizon. Usually for chronic illnesses (i.e. NAFLD), a lifetime time horizon is recommended.(137)

4.4.5 Discounting

Discounting is an important concept that revolves around the notion that costs and outcomes today are more valuable than costs and outcomes that occur in the distant future. It reflects the universal time preference for benefits earlier rather than later. It also takes into account the opportunity cost of capital or the return on investment that could have been gained if the investment of resources were elsewhere. Therefore, costs and outcomes should be discounted relative to their present value. In health economic analyses, discount rates are usually set at 3-5% and reflects government or market interest rates for cost of capital whose maturity is over a long time horizon.(137) In recent years in Canada, the discount rate has been recommended at 1.5%, which better reflects long-term borrowing costs for Canadian provinces.(137)

The basic formula for calculating present values for a given discount rate is shown below. For example, $1000 presently would in 5 years, at a discount rate of 3%, yield a value of $888.

$$P = \sum_{n=1}^{n} \frac{Fn}{(1+r)^n}$$

P = present value
F = future cost (or benefits) at year n
r = annual discount rate

4.4.6 Utilities (QALYs)

Health-related quality of life (HRQoL) is a multifaceted measure of the effect of an intervention or disease process on an individual's overall well-being.(137) This measure encompasses physical and psychological well-being, social and occupational functioning, and somatic sensation. There are many methods to measure HRQoL. For the purposes of cost-utility analyses, preference-based measures are used because they provide an overall numerical value for a given HRQoL state.(137)

The strict definition of "utilities" refers to preferences obtained by methods that involve choice made under uncertainty.(137) Broadly speaking, the term "utilities" can be used to describe the numeric weight assigned to quantify a particular HRQoL state. Utilities are based on an individual's preference for better health states. Thus, a utility of 1 represents perfect health compared with a utility of 0 which represents a health state equivalent to death.(137)

Quality-adjusted life years or QALYs is the recommended outcome in cost-utility analyses. It is superior to simple life-years gained as an outcome measure because it provides information on both quantity and quality of life lived.(140) One of the advantages of QALYs is that it can be used to standardize and compare multiple different and unrelated conditions.

Health-related quality of life can be measured using various tools. These include direct methods of standard gamble, time trade off, and rating scales. In practice, indirect hybrid tools such as the Health Utility Index questionnaire and the EuroQol instrument (EQ-5D) questionnaire are often used. These questionnaires incorporate many essential dimensions of living (i.e. vision, hearing, speech, ambulation, emotion etc.) in order to measure general health status. (141-143)

4.4.7 Cost-effectiveness criteria

For an understanding of the basic approach to cost-effectiveness, refer to **Figure 4.4** below. Figure 4.4 illustrates a cost-effectiveness plane whereby the *center (0,0)* represents the current standard of care, the *x* axis represents effectiveness in QALYs, and the *y* axis represents costs in $. Any new interventions may theoretically have higher or lower effectiveness and higher or lower costs. If an intervention has both higher costs and lower

effectiveness than its reference, it is deemed *dominated* and thus rejected as possible replacement strategy. Conversely, if an intervention has both lower costs and higher effectiveness than its reference, it is considered *dominant* and more cost-effective. In both these scenarios, no further analysis is needed.(29, 144)

However, if an intervention falls in the other two quadrants on the cost-effectiveness plane, where it either has higher effectiveness and costs or lower effectiveness and costs than the reference, then further analysis is required. In these situations, analysis of the incremental gains in effectiveness and costs are needed.(29, 144)

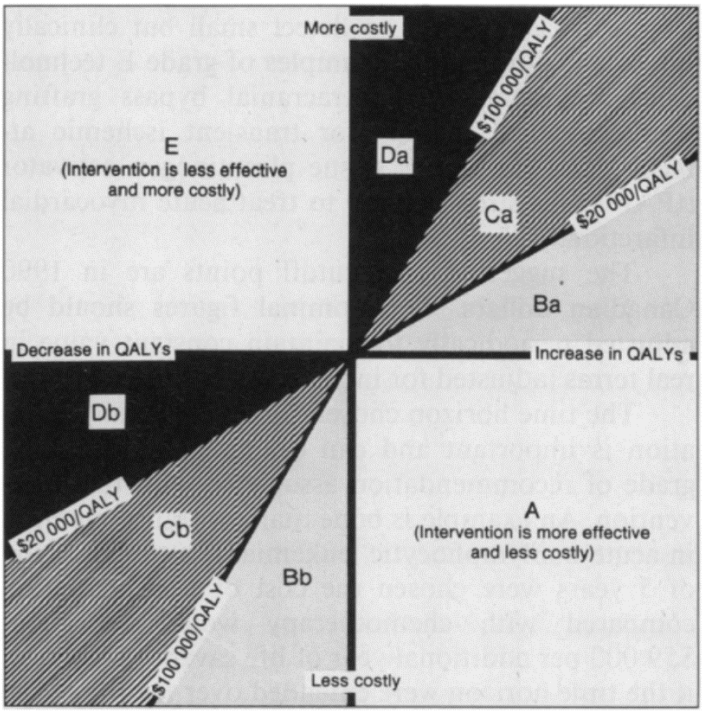


Figure 4.4: Cost-effectiveness graph. Grade A interventions are less costly and more effective and should be chosen in an ideal situation. Grade E interventions are dominated in terms of both effectiveness and price, and thus should not be introduced. Interventions in the right upper quadrant are more effective and more costly than their alternatives. Interventions in the left lower quadrant are less effective and less costly. The decision to implement either of these interventions depend on their incremental cost-effectiveness ratios and willingness-to-pay threshold. Laupacis A *et al.* How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *CMAJ : Canadian Medical Association journal*. 1992;146(4):473-81. (145)

4.4.8 Incremental cost-effectiveness ratio

If an intervention is not *dominant* or *dominated* by its comparator, incremental cost-effectiveness ratios (ICER) are important outcome measures in both cost-utility and cost-effectiveness analyses. ICER is defined as the difference between the costs of two possible interventions divided by the difference in their outcome/effect. In the case of cost-utility analyses, the cost would be in dollars and the outcome would be QALYs (i.e. $/additional QALY gained). In the case of cost-effectiveness analyses, the cost would be in dollar and the outcome would be in natural health units (i.e. $/additional life-year gained). The ICER allows for analysis between the incremental gains of one intervention versus another. (137)

$$ICER = \frac{(C_1 - C_0)}{(E_1 - E_0)}$$

4.4.9 Willingness-to-pay threhold

The willingness-to-pay threshold is a predefined cost-effectiveness ratio that helps guide researchers in deciding cost-effectiveness when looking at different ICERs. It is usually used in cost-utility studies where thresholds are set in $/QALY. Put differently, it can be thought of as the price per additional QALY gained. Historically, in cost-utility analysis, the price of an additional QALY gained has been set at $50,000/QALY. This number is attributed historically to cost-effectiveness ratio for dialysis in the 1970s, for which the government was content to pay.(146) However, more recent literature suggests that this number is too low and that the willingness-to-pay threshold should be set as high as $100,000 to $200,000/QALY.(147)

4.5. Modelling in health economics

4.5.1 Decision analysis

Decision analyses comprise an explicit way of addressing specific clinical decisions by deconstructing complex clinical scenarios in logical decision steps known as a decision tree. A decision tree illustrates all plausible alternatives, relationships, and outcomes in a specific

clinical scenario.(29, 148) At each step of the decision tree, there are corresponding probability and outcome values, which are derived from literature. By incorporating these different probability and outcome values, the decision analysis can come to defined conclusions based on the average expected result.(29) Using the decision tree, the decision maker can then weigh the different options in a more accurate and objective way, ultimately facilitating a more informed clinical decision.(29, 144, 148) **Figure 4.5** provides an overview of steps involved in decision analysis.

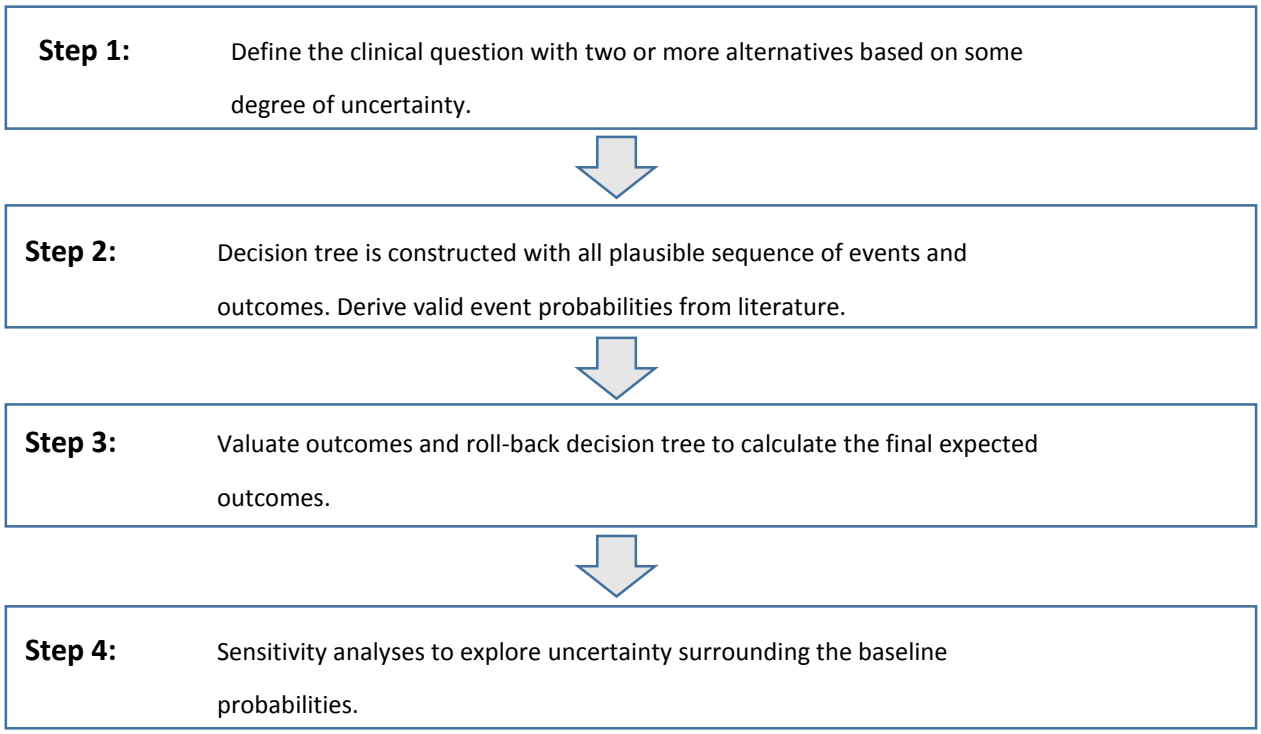


Figure 4.5: Stepwise approach to decision analysis. Adapted from: Briggs AC, K.; Sculpher M. . Decision modelling for health economic evaluation. 1st ed. USA: Oxford University Press, Inc. ; 2006. 224 p.(144)

4.5.1 Building a decision tree

There are three main steps in creating a decision analysis tree. Firstly, the clinical problem should be well defined and structured. This includes an exhaustive list of all possible comparators in a specific clinical question as well as all relevant outcome measures.(29, 144)

For example **Figure 4.6**, in the detection of a generic illness, there are two possible alternatives: modality A and modality B.

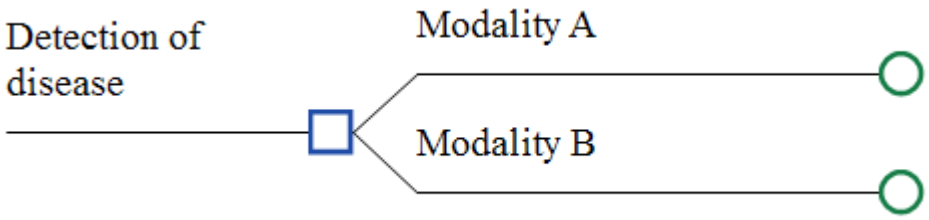


Figure 4.6: Beginning a decision tree. The hypothetical clinical question is choice between two modalities for the detection of a disease. The blue square denotes the choice between two mutually exclusive modalities A and B.

Next, a decision tree is constructed based on all plausible and logical sequence of events and outcomes associated with each alternative. Within a decision tree, there are two major categories of events: those that require a decision and those that are probabilistic in nature. The former is indicated on the decision tree by convention as a square, whereas the latter is denoted by convention as a circle. The probability of each step is found or estimated according to the best data in literature. Sources for these probabilities include randomized-controlled trials, meta-analyses, prospective and retrospective studies, administrative data banks, and expert opinions. The stronger the evidence, the more valid the decision model.(29, 144) Refer to **Figure 4.7** as an example of an expanded decision tree with probability values, assuming each modality carries a probability of adverse side effect. By convention, the end of a branch in a decision tree is denoted by a red triangle.

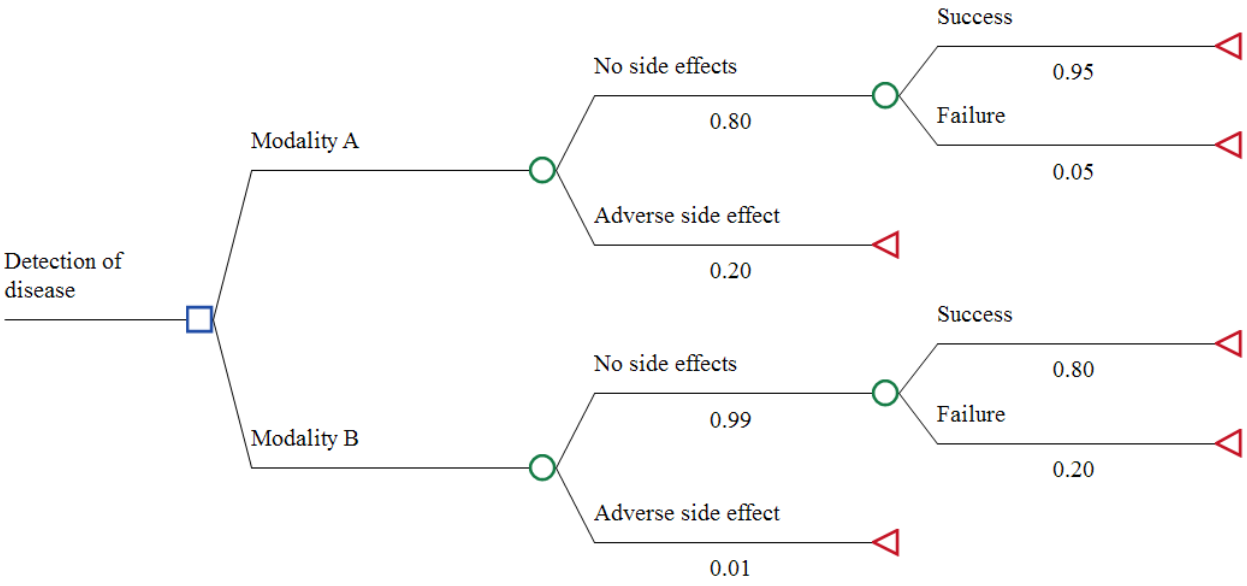
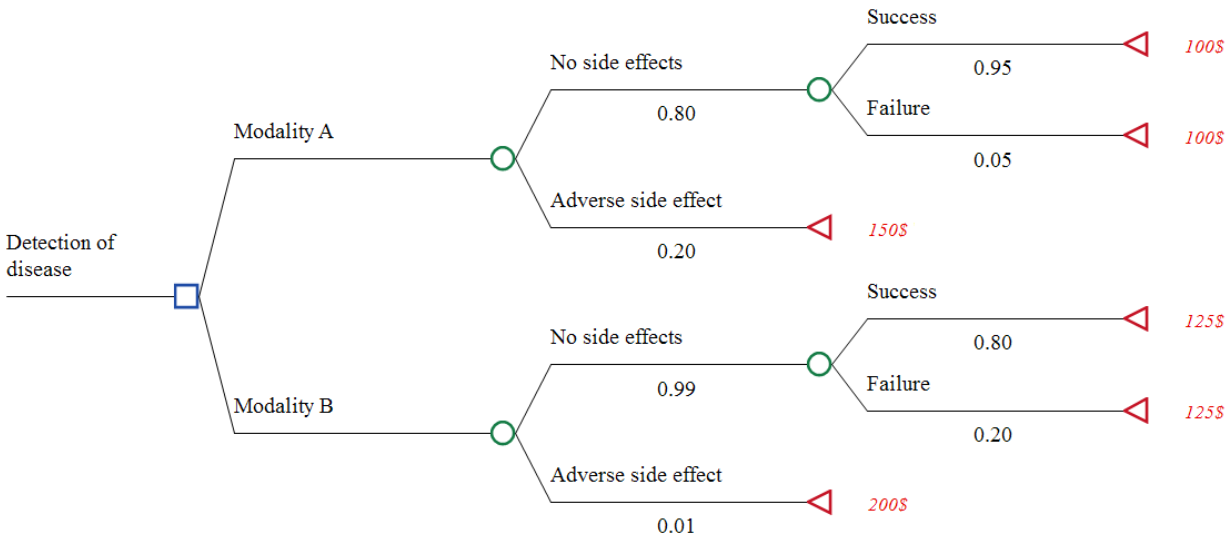


Figure 4.7: Example of an expanded decision tree with step probabilities. The green circle indicates probabilistic nodes. The red triangle indicates the end of a decision branch. The probability of each event is derived from literature. The validity of a decision tree depends on the validity of the data sources.

The third step is to find and valuate outcome measures for each event in the decision tree (**Figure 4.8**). These outcome measures can be related to dollar costs or non-monetary values such as life years or quality-adjusted life years. At this point, it is possible to calculate or "roll-back" the expected outcomes in each decision. This is done by multiplying the respective probability values of each decision branch with the respective outcome values, and calculating the product-sum of each intervention.(29, 144) Thus, the decision tree provides its conclusion in terms of an average expected result, as denoted by monetary and non-monetary values, which will ultimately help in the clinical decision making process.(29, 144)



Average expected cost of modality A = (0.80*0.95*$100)+(0.80*0.05*$100)+(0.20*$150)= $110
Average expected cost of modality B = (0.99*0.80*$125)+(0.99*0.20*$125)+(0.01*$200)= $125.75

Figure 4.8: Example of outcome calculation is a decision tree. In this example, a dollar cost is assigned to each branch of the decision tree. By calculating the product-sum of each intervention, the decision-maker arrives at the conclusion that modality A is on average less expensive than modality B for detection of a generic disease.

4.5.3 Sensitivity analysis

Sensitivity analyses are important tools for accounting for uncertainty and assessing the level of confidence associated with a conclusion in an economic analysis. Any estimate of probabilities, outcomes, and costs are subject to uncertainty from literature discrepancies, biologic variation, and differing techniques. Sensitivity analysis is performed by adjusting different key assumptions in an evaluation and recording the impact of these variations on the result of the evaluation. One or more variables can be changed while holding other variables constant, which allows for detection of the most important variables on the final outcome. Sensitivity analyses also can be a useful method for detecting errors within a decision tree as a "de-bugging" tool.(29, 144)

There are two main types of sensitivity analyses. In deterministic sensitivity analysis, the model input is specified as multiple point estimates and varied manually. In probabilistic sensitivity analysis, model inputs are specified as a distribution and varied. In recent years, the CADTH has recommended probabilistic sensitivity analyses as the reference standard for economic analyses, which has less potential to have biases estimates of costs and outcomes.(136)

4.5.4 Markov models

4.5.4.1 Introduction to Markov models

Markov models are repetitive decision trees that are used to simulate predictable events that may occur repeatedly over a period of time. They are useful for modeling clinical situations where there is a risk that is ongoing over time. In the clinical setting, appropriate examples for Markov modeling would include, for example, ongoing risk of rupture of abdominal aortic aneurysms. Other appropriate examples would involve screening for disease at fixed intervals over a long period of time, for example, colorectal screening programs.(30) The main advantage of Markov modelling is that it explicitly accounts for the passage of time in the model, which, by comparison, is a factor that is not explicitly accounted for in standard decision trees. It allows for more complex modelling, more in line with real-life clinical scenarios.

4.5.4.2 Markov states

In a Markov model, the theoretical patient is assumed to always be in one of a finite number of clinical health states, which is termed *Markov states*.(30) Markov states are predefined and, in the clinical setting, usually is based around the natural history of a chronic disease process. For example, the progression of chronic liver disease from healthy, to steatosis, to steatohepatitis, and fibrosis (**Figure 4.9**). Death usually is omnipresent as a background Markov state.

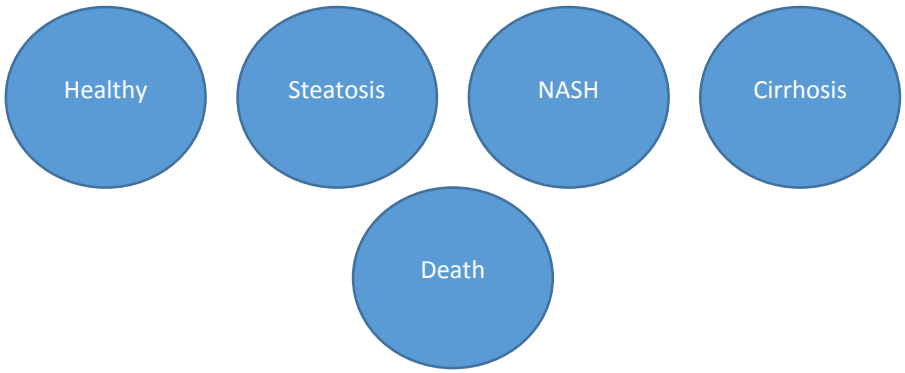


Figure 4.9: Markov states in NAFLD. In the clinical setting, Markov states are usually determined based on understanding of natural history.

4.5.4.3 Markov cycles and transitions

Markov states interact with one and another through transition probability. The modeled patient can transition from one state to another via predefined routes, which are in accordance with established natural or clinical history of the process at hand. The time horizon of a Markov model is divided into predefined and equal increments of time, known as *Markov cycles*.(30) During each cycle, the patient has a transition probability of progressing or remaining from one state to another. Markov cycles are usually selected based on relevance to the clinical question. For example, for chronic disease process, an appropriate Markov cycle would be every 1 or 2 years. On the other hand, for more acute processes where the rate of change is higher, the Markov cycle may be monthly, weekly, or even daily.(30) In the example of chronic liver disease, the established transition from steatosis to cirrhosis is well documented in literature (**Figure 4.10**).

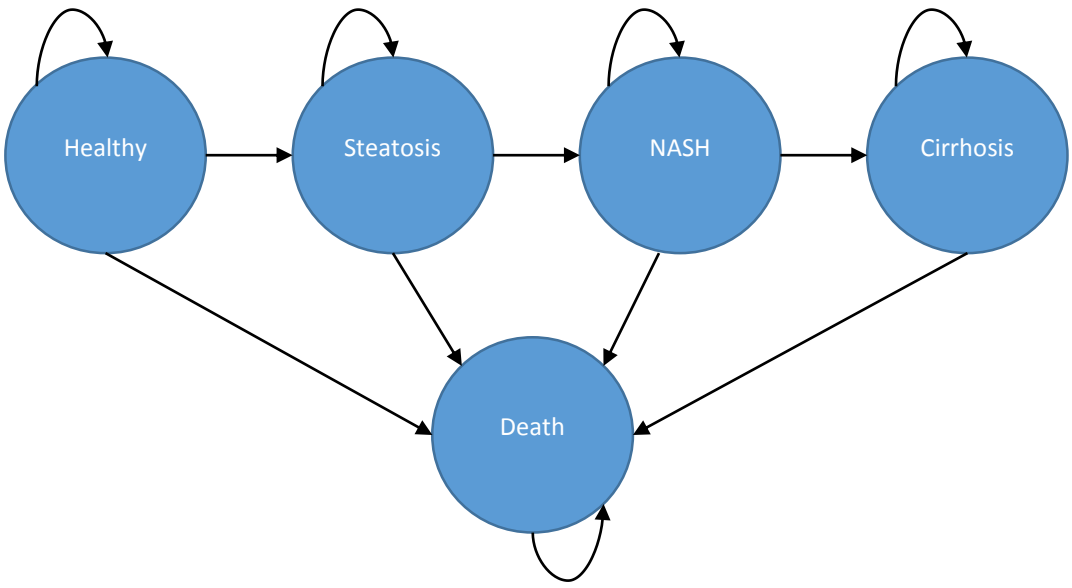


Figure 4.10: Markov states of NAFLD in transition. The Markov cycle in this case may be longer term (i.e. yearly) given that chronic liver disease develops over the long term. Arrows that lead from one state to another indicate transition during a cycle. Arrows that lead from one state to itself indicate the theoretical patient will remain in that state for a cycle. The death state is an *absorbing state* in the sense that patients who enter will remain until the termination of the model. It is important to note that if enough Markov cycles have passed, all of the theoretical cohort of patients will enter into the absorbing state.(30)

4.5.4.4 Calculating outcomes/utilities

In Markov models, one or multiple outcome values are assigned to each Markov state. These outcome measures may be simple, such as life years spent in each health state or more complicated, such as health-related quality of life associated with each health state. Results are calculated based on the average time spent in each Markov state. For example, if the outcome measure of chronic liver disease is survival duration in a specific state, then one would have to add together the average time spent in that state.(30)

$$\text{Expected utility} = \sum_{s=1}^{n} t_s$$

t_s is time spent in state **s**.

On the other hand, if the outcome measure is quality-adjusted life years, then one would have to take into account and multiple the quality-adjusted life year for that particular state by the total life year spent.(30) In the specific example of cost-effectiveness studies, a specific utility and cost is assigned to each Markov state and the model is evaluated separately for cost and utility. The incremental cost-effectiveness ratios are calculated in the same manner as in standard decision trees. (30)

$$\text{Expected utility} = \sum_{s=1}^{n} t_s \times u_s$$

$\mathbf{t_s}$ is time spent in state **s**.

$\mathbf{u_s}$ is the specific utility (QALY) of the state **s**.

4.5.4.5 The Markov property

In Markov modelling, the underlying property or restriction is that "the behavior of the process subsequent to any cycle depends only on its description in that cycle."(30) Put differently, a patient entering from one state to another has no memory of prior cycles. In the example of chronic liver disease, a cirrhotic patient will enter into the death state depending on a specific probability. This probability will not change depending on how long that patient has spent in the healthy, steatotic, or NASH states prior to entering in the cirrhotic state. Thus, in order to accurately model the overall prognosis of a patient, a distinct and separate state should be created for each subset of the population that has a different prognosis.(30)

Chapter 5
Research project

5.1 Overview of problem

Nonalcoholic fatty liver disease (NAFLD) has become the most common cause of chronic liver disease in the Western world due to improved management of viral hepatitis and the rising obesity and diabetes epidemic.(129) Healthcare costs are estimated to be significant.(10) They are anticipated to rise, with nonalcoholic steatohepatitis (NASH) anticipated to become the leading cause of liver transplantation by 2020.(12)

The current gold standard for diagnosis is liver biopsy, a method that is characterized by significant rates of morbidity and mortality.(101) In response, new biomarker and imaging modalities have been developed to answer the need for non-invasive measures.(92) Early detection would facilitate the implementation of a multidisciplinary treatment algorithm, which involves both lifestyle changes and pharmacologic therapy, as recommended by current guidelines.(1, 5)

However, the current healthcare environment in Canada is characterized by mounting costs and increasing resource scarcity. The same applies for the majority of developed nations, where a large percentage of the gross domestic product (GDP) is utilized in healthcare.(136) In this resource-scarce setting, any new screening program should be analyzed in terms of cost-effectiveness.

There is a current need to investigate the cost-effectiveness of a screening program for NAFLD.(78) Prior to this study, no other cost-effectiveness study has been undertaken.

5.2 Hypothesis and objectives

Primary hypothesis: Screening of NAFLD in the general population will be cost-effective. Furthermore, screening of NAFLD in high-risk groups (obesity and type 2 diabetes) would be even more cost-effective.

Research objectives:

I. To perform a cost-utility analysis of screening for NASH using the most validated non-invasive methods.

II. To identify key factors that drive cost-effectiveness in order to prioritize areas for future research.

5.3 Significance of project

This project will be the first of its kind in addressing a key knowledge gap in the study of NAFLD. If screening is found to be cost-effective in the general population and/or in high-risk groups, this can potentially have real consequences, laying the foundation for a new screening program in Canada. In the literature review phase of the study, we will highlight current deficiencies in our knowledge of NAFLD. During the sensitivity analysis of our model, we will identify key drivers of cost-effectiveness.

Chapter 6
Scientific contribution

Chapter 7
Conclusion

7.1 Summary

Non-alcoholic fatty liver disease (NAFLD), which encompasses a spectrum of disease ranging from simple steatosis, to non-alcoholic steatohepatitis (NASH), to end-stage liver cirrhosis, is currently the leading cause of chronic liver disease in the developed world. The costs associated with this disease are enormous, both to the healthcare system and to society as a whole. In this thesis, we addressed and dissected this complex disease of growing importance.

We performed a systematic analysis of this topic, looking at the epidemiology and natural history of this continuum of disease. We examined the current non-invasive diagnostic modalities available for NAFLD, including both biomarker panels and imaging-based methods, comparing them with liver biopsy, the current reference standard for diagnosis. Finally, we looked at the current management recommendations for NAFLD, which currently encompasses lifestyle changes and pharmacologic adjuncts.

Before a health economic analysis was performed, we analyzed the appropriateness of screening for NAFLD, as outlined by Wilson and Jungner at the World Health Organization. These appropriateness criteria guided the methodology of our cost-utility study. For example, a main principle in preventing over-diagnosis is to differentiate between diseases with a benign natural history and those disease processes that progress to cause more harm. Thus, in our cost-utility model, we did not choose to screen for simple steatosis since literature suggests that it is not likely to develop into advanced liver disease over a lifetime.

Finally, we examined economic analysis in healthcare and the different types of economic analyses that are performed. A cost-utility study was deemed most appropriate, given it addresses both cost-effectiveness and takes into account the quality of life of patients at different stages of this complex disease.

7.2. Creation of the TreeAge model

To our knowledge, this is the first study addressing the cost-effectiveness of NASH screening in both a general population and in high-risk obese or diabetic populations. When determining our screening strategies, we built on the strategies proposed by Musso et al. in their meta-analysis of

the natural history of NAFLD and the current noninvasive strategies.(9) The screening strategies proposed comprise of noninvasive measurements of NASH and NASH-related fibrosis. The question of whether to screen for steatosis arose due to several reasons. First, it represents a substantial percentage of the population, and much larger than the subset affected by NASH or NASH-related fibrosis.(129) Second, there are successful lifestyle modifications that can halt and potentially reverse the disease process.(213) However, as discussed by Musso et al., screening for simple steatosis would not be cost-effective. In our current understanding of NAFLD, simple steatosis is generally considered a benign disease with a small proportion progressing to NASH.(56) By itself, steatosis has no significant bearing on one's health-related quality of life.(171) Furthermore, unlike NASH or NASH-related, simple steatosis by itself is not related to increased liver-related mortality.(55, 56) Put together, screening for steatosis would necessarily be ineffective and costly. All screening strategies which include screening for steatosis at an early stage would be economically dominated by those strategies that focus on more severe stages of NAFLD such as NASH or NASH-fibrosis, which comprises significant increases to liver-related mortality.(53)

Figure 7.1 illustrates a conceptual example of the decision tree used in our simulations regarding the natural history of disease. In **Figure 7.1**, the natural history of NAFLD is best approximated according to current understandings. The model assumes an annual cycle length, meaning that each completion through the decision tree occurs within one year's time. We can appreciate that the progression from well to steatosis occurs according to an annual transitional probability as denoted by "*t_WellSteatosis*". Before such a transition occurs, the probability of annual all-cause mortality is taken into account, which is related to the patient's age as derived from Statistics Canada. Once a transition (to the steatosis category) or non-transition (remaining in the well category) has been established, there is an assigned cost, i.e. "*c_Nocare*", and utility value, i.e. "*e_Steatosis*". For the purposes of discussion and conceptual illustration, this example is but a subset of a much larger decision tree. As explained in the research paper, all data, including annual transitional probabilities, costs, and utilities are derived from our extensive literature review.

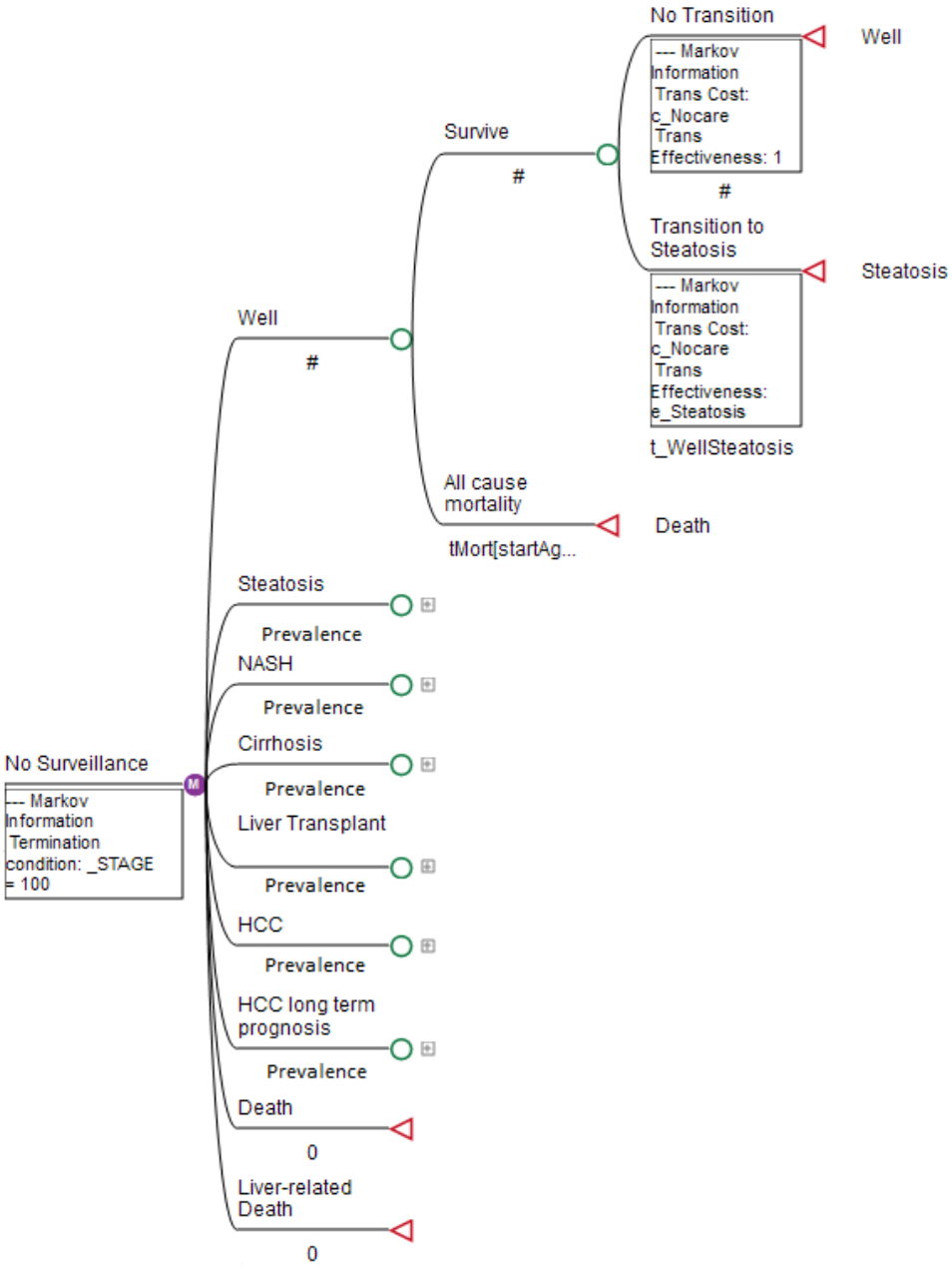


Figure 7.1: Schematic tree diagram of the natural history of NAFLD. *"_STAGE = 100"* denotes the number of years the Markov model will run for. In this case, the model ends after 100 simulated years. For the purposes of the study, 100 years was an arbitrary number selected to ensure a lifetime horizon (i.e. the simulation ends when all patients have died).

The integration of both screening and treatment algorithms presented its own challenges. **Figure 7.2** illustrates conceptually how screening was integrated into the life cycle. Please note that this example is a subset of the entire screening strategy entitled "*NFS/ARFI/CK18 + Biopsy confirmation + Vit E*". This algorithm is illustrated in **Figure 2B** of the research paper. We note that as the simulated patient progresses in the "*NASH*" disease state, he/she is given an annual transitional probability of remaining in the "*NASH*" disease state, progressing to the liver "*F1 fibrosis*" disease state, or dying of either all-cause or liver-related mortality. However, within the natural history of progression, we insert the underlying screening program. In this example, those who remain in the "NASH" disease state are assessed with NAFLD fibrosis score and the CK-18 blood test. According to the sensitivity probabilities of each test, the simulated patient is diagnosed with the disease or not. Those found to have NASH through the CK-18 blood test is confirmed with liver biopsy, which assumes a perfect sensitivity. The underlying costs of each test as well as the utility value of the disease state is assigned at the end of the decision branch. Patients with a diagnosis of NASH progress to the "*NASH Treatment*" decision branch.

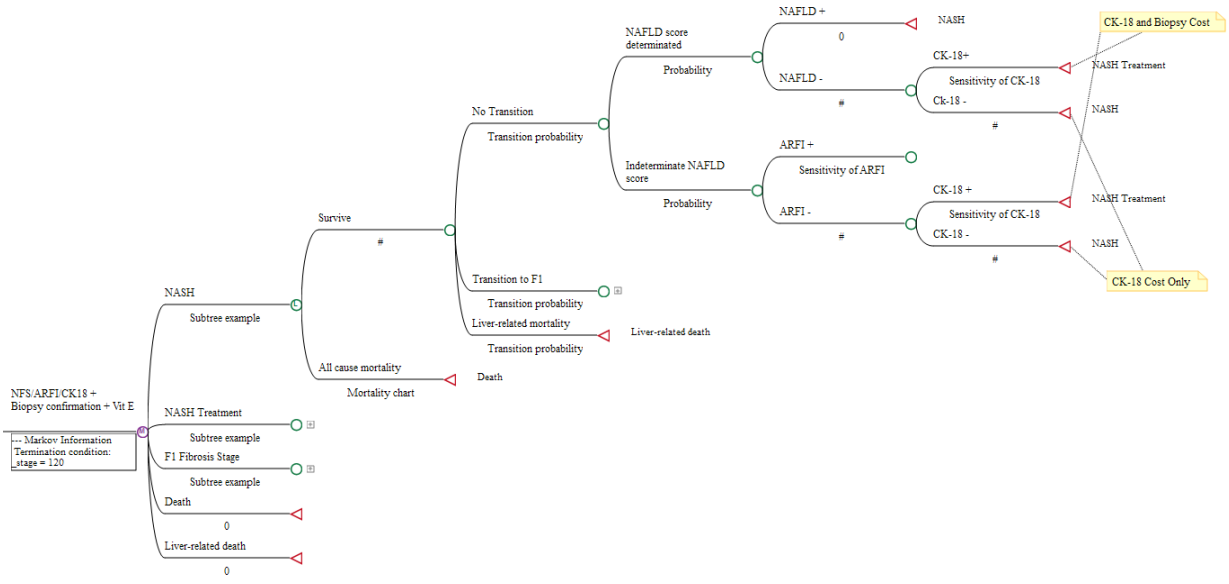


Figure 7.2: Integration of screening strategies into natural history. For the purposes of illustration, "*NASH Treatment*" and "*F1 Fibrosis Stage*" decision branches are collapsed.

Integration of treatment in the decision tree follows the same line of thought. **Figure 7.3** illustrates conceptually a subset of the entire tree for the screening strategy entitled "*NFS/ARFI/CK18 + Biopsy confirmation + Vit E*". We note that the NASH diagnosed patients who survive all-cause mortality are then treated according to the screening strategy, in this case vitamin E. Compliance and non-compliance rates are taken into account for the simulation. Those that are treated with vitamin E have a probability to stay stable for the year, transition to early liver fibrosis, or die of liver-related mortality. The same options are present for patients who are non-compliant, but with different probabilities.

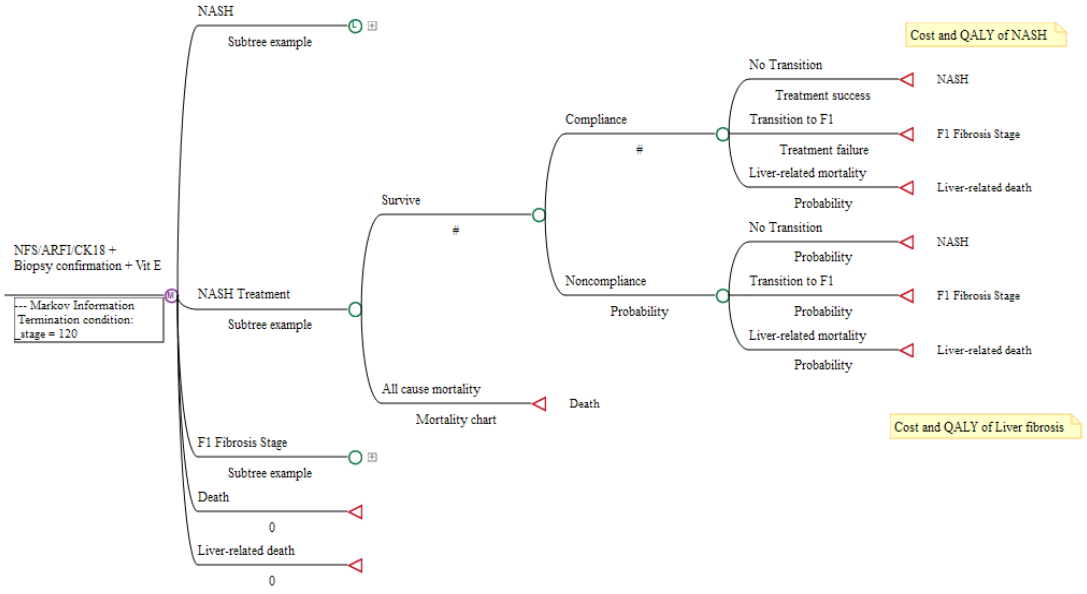


Figure 7.3: Integration of treatment strategies. For the purposes of illustration, "*NASH*" and "*F1 Fibrosis Stage*" decision branches are collapsed.

7.3. Discussion

The above examples illustrate the conceptual thinking behind the decision tree used in the research project. Once a screening strategy was fully mapped, other screening strategies were built

in similar fashion and ultimately compared in cost-utility analyses and sensitivity analyses. While there is no reference standard method of building Markov models, there are a few theoretical advantages for this model method used. This model excels at illustrating and approximating the current natural history of NAFLD. The minutiae of transitional probability, mortality rates, costs, health-related quality of life, and treatment effects are all accounted for.

However, there are also a few theoretical limitations to this model. Since the natural history of the disease forms the foundation for the decision tree, all patients are categorized into either having a disease or not having a disease. The true negative rate or the specificity of the diagnostic examinations is not explicitly accounted for in this model. A false-positive study would lead to over-diagnosis and over-treatment, and therefore not having this option presents biases in terms of overall cost-effectiveness of strategies. Future models should attempt to address this limitation.

Besides intrinsic model limitations, there are several additional points that would be interesting to address in future research. Firstly, variations in screening interval were not tested in the study because we made the assumption that screening would be undertaken in yearly interval. This is based on precedence from prior cost-effectiveness analysis related to treatment options for NASH and NASH-fibrosis.(155) Nonetheless, current acceptable screening practices in breast and colon cancer do not necessary revolve around an annual interval. Secondly, since the publication of the study, new Canadian guidelines for the economic evaluation of health technologies recommend the use of probabilistic sensitivity analyses over deterministic sensitivity analyses due to less overall bias and more robustness of the analyses.(137) Lastly, given that current diagnosis of NAFLD requires liver biopsies, we did not explicitly account for the decreased number of unnecessary liver biopsies in the diagnosis of NAFLD and NASH in the setting of NASH screening versus no surveillance.

7.3. Future direction

Given our current knowledge, we found that screening is cost-effective, particularly in high-risk populations with metabolic syndrome. This work could potentially have important implications for patients and their families, healthcare providers, as well as healthcare systems. A screening program for NAFLD would initiate early recognition and treatment of disease, allowing

patients to address the disease at a preclinical phase. Healthcare providers would also place increasing importance on screening and prevention. Over the long term, this would have cost-saving effects for healthcare systems, which is important, given the current need to addressing increasing healthcare spending.

At the same time, our research outlined current knowledge gaps concerning NAFLD that will require future research and direction. The bulk of our understanding of the natural history of NAFLD comes from relatively short-term observational studies. Given the chronicity of the disease, controlled, longitudinal studies over the long-term are still needed to definitively understand the natural history. Furthermore, there are still relevant knowledge gaps concerning the pharmacologic treatment of NASH and NAFLD. These include addressing questions about the long-term efficacy and adverse effects of different pharmacologic therapy, in the form of long-term multicenter randomized controlled trials. Finally, specific health-related quality of life studies in NASH should be undertaken.

All new noninvasive examinations currently and necessarily use liver biopsy as the reference standard. It is inherently difficult to assess the sensitivity and specificity of liver biopsy because besides hepatectomy or whole liver explantation, there is no better standard available. However, as mentioned previously, liver biopsy is an imperfect reference and limited by errors relating to sample location and size. By assuming the accuracy of liver biopsy to be perfect, we are overestimating the diagnostic ability of liver biopsy and therefore underestimating the number of missed diagnoses/treatments. That being said, until a better reference standard is established, liver biopsy remains the current reference standard for diagnosis of NAFLD spectrum.

This study represents a first step in determining the clinical paradigm surrounding NAFLD and will act as a stepping stone for future economic analyses.

Bibliography

1. Ratziu V, Bellentani S, Cortez-Pinto H, Day C, Marchesini G. A position statement on NAFLD/NASH based on the EASL 2009 special conference. Journal of hepatology. 2010;53(2):372-84.
2. Neuschwander-Tetri BA, Caldwell SH. Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. Hepatology. 2003;37(5):1202-19.
3. Medina J, Fernandez-Salazar LI, Garcia-Buey L, Moreno-Otero R. Approach to the pathogenesis and treatment of nonalcoholic steatohepatitis. Diabetes care. 2004;27(8):2057-66.
4. Bhala N, Jouness RI, Bugianesi E. Epidemiology and natural history of patients with NAFLD. Curr Pharm Des. 2013;19(29):5169-76.
5. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. Gastroenterology. 2012;142(7):1592-609.
6. Wilson JMG JGPaposfdG, Switzerland: World Health Organization; 1968. Report No.: Public Health Papers No. 34. Available from: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf.
7. Fernandez-Salazar L, Velayos B, Aller R, Lozano F, Garrote JA, Gonzalez JM. Percutaneous liver biopsy: patients' point of view. Scandinavian journal of gastroenterology. 2011;46(6):727-31.
8. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005;128(7):1898-906.
9. Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. Annals of medicine. 2011;43(8):617-49.
10. Baumeister SE, Volzke H, Marschall P, John U, Schmidt CO, Flessa S, et al. Impact of fatty liver disease on health care utilization and costs in a general population: a 5-year observation. Gastroenterology. 2008;134(1):85-94.
11. Kim WR, Brown RS, Jr., Terrault NA, El-Serag H. Burden of liver disease in the United States: summary of a workshop. Hepatology. 2002;36(1):227-42.
12. Charlton MR, Burns JM, Pedersen RA, Watt KD, Heimbach JK, Dierkhising RA. Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. Gastroenterology. 2011;141(4):1249-53.
13. Xu J, Kochanek KD, Murphy SL, Tejada-Vera B. Deaths: final data for 2007. Natl Vital Stat Rep. 2010;58(19):1-19.
14. Younossi ZM, Stepanova M, Afendy M, Fang Y, Younossi Y, Mir H, et al. Changes in the prevalence of the most common causes of chronic liver diseases in the United States from 1988 to 2008. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2011;9(6):524-30 e1; quiz e60.
15. Yang L, Colditz GA. Prevalence of Overweight and Obesity in the United States, 2007-2012. JAMA Intern Med. 2015.

16. Gotay CC, Katzmarzyk PT, Janssen I, Dawson MY, Aminoltejari K, Bartley NL. Updating the Canadian obesity maps: an epidemic in progress. Can J Public Health. 2013;104(1):e64-8.
17. Farrell GC, Larter CZ. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. Hepatology. 2006;43(2 Suppl 1):S99-S112.
18. Charlton M. Nonalcoholic fatty liver disease: a review of current understanding and future impact. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2004;2(12):1048-58.
19. Loomba R, Sanyal AJ. The global NAFLD epidemic. Nat Rev Gastroenterol Hepatol. 2013;10(11):686-90.
20. Farrell GC, Wong VW, Chitturi S. NAFLD in Asia--as common and important as in the West. Nat Rev Gastroenterol Hepatol. 2013;10(5):307-18.
21. Fan JG, Jia JD, Li YM, Wang BY, Lu LG, Shi JP, et al. Guidelines for the diagnosis and management of nonalcoholic fatty liver disease: update 2010: (published in Chinese on Chinese Journal of Hepatology 2010; 18:163-166). J Dig Dis. 2011;12(1):38-44.
22. Das K, Das K, Mukherjee PS, Ghosh A, Ghosh S, Mridha AR, et al. Nonobese population in a developing country has a high prevalence of nonalcoholic fatty liver and significant liver disease. Hepatology. 2010;51(5):1593-602.
23. Singh SP, Nayak S, Swain M, Rout N, Mallik RN, Agrawal O, et al. Prevalence of nonalcoholic fatty liver disease in coastal eastern India: a preliminary ultrasonographic survey. Trop Gastroenterol. 2004;25(2):76-9.
24. Amarapurkar D, Kamani P, Patel N, Gupte P, Kumar P, Agal S, et al. Prevalence of non-alcoholic fatty liver disease: population based study. Ann Hepatol. 2007;6(3):161-3.
25. Chitturi S, Farrell GC, Hashimoto E, Saibara T, Lau GK, Sollano JD, et al. Non-alcoholic fatty liver disease in the Asia-Pacific region: definitions and overview of proposed guidelines. Journal of gastroenterology and hepatology. 2007;22(6):778-87.
26. Mendez-Sanchez N. Non alcoholic fatty liver disease. Ann Hepatol. 2009;8 Suppl 1:S3.
27. Bodenheimer T, Chen E, Bennett HD. Confronting the growing burden of chronic disease: can the U.S. health care workforce do the job? Health affairs. 2009;28(1):64-74.
28. Giacomini M. How good is good enough? Standards in policy decisions to cover new health technologies. Healthcare policy = Politiques de sante. 2007;3(2):91-101.
29. Petitti DB. Meta-analysis, decision analysis, and cost-effectiveness analysis: methods for quantitative synthesis in medicine. 2nd ed. New York: Oxford University Press, Inc. ; 2000. 305 p.
30. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Medical decision making : an international journal of the Society for Medical Decision Making. 1993;13(4):322-38.
31. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology. 1999;116(6):1413-9.
32. Adams LA, Sanderson S, Lindor KD, Angulo P. The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. Journal of hepatology. 2005;42(1):132-8.
33. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Alimentary pharmacology & therapeutics. 2011;34(3):274-85.

34. Marcos A, Fisher RA, Ham JM, Olzinski AT, Shiffman ML, Sanyal AJ, et al. Selection and outcome of living donors for adult to adult right lobe transplantation. Transplantation. 2000;69(11):2410-5.
35. Lee JY, Kim KM, Lee SG, Yu E, Lim YS, Lee HC, et al. Prevalence and risk factors of non-alcoholic fatty liver disease in potential living liver donors in Korea: a review of 589 consecutive liver biopsies in a single center. Journal of hepatology. 2007;47(2):239-44.
36. Caballeria L, Pera G, Auladell MA, Toran P, Munoz L, Miranda D, et al. Prevalence and factors associated with the presence of nonalcoholic fatty liver disease in an adult population in Spain. European journal of gastroenterology & hepatology. 2010;22(1):24-32.
37. Bedogni G, Miglioli L, Masutti F, Tiribelli C, Marchesini G, Bellentani S. Prevalence of and risk factors for nonalcoholic fatty liver disease: the Dionysos nutrition and liver study. Hepatology. 2005;42(1):44-52.
38. Williams CD, Stengel J, Asike MI, Torres DM, Shaw J, Contreras M, et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. Gastroenterology. 2011;140(1):124-31.
39. Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatology. 2004;40(6):1387-95.
40. Patt CH, Yoo HY, Dibadj K, Flynn J, Thuluvath PJ. Prevalence of transaminase abnormalities in asymptomatic, healthy subjects participating in an executive health-screening program. Digestive diseases and sciences. 2003;48(4):797-801.
41. Clark JM, Brancati FL, Diehl AM. The prevalence and etiology of elevated aminotransferase levels in the United States. The American journal of gastroenterology. 2003;98(5):960-7.
42. Boza C, Riquelme A, Ibanez L, Duarte I, Norero E, Viviani P, et al. Predictors of nonalcoholic steatohepatitis (NASH) in obese patients undergoing gastric bypass. Obesity surgery. 2005;15(8):1148-53.
43. Machado M, Marques-Vidal P, Cortez-Pinto H. Hepatic histology in obese patients undergoing bariatric surgery. Journal of hepatology. 2006;45(4):600-6.
44. Prashanth M, Ganesh HK, Vima MV, John M, Bandgar T, Joshi SR, et al. Prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus. J Assoc Physicians India. 2009;57:205-10.
45. Leite NC, Salles GF, Araujo AL, Villela-Nogueira CA, Cardoso CR. Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. Liver international : official journal of the International Association for the Study of the Liver. 2009;29(1):113-9.
46. Chen ZW, Chen LY, Dai HL, Chen JH, Fang LZ. Relationship between alanine aminotransferase levels and metabolic syndrome in nonalcoholic fatty liver disease. J Zhejiang Univ Sci B. 2008;9(8):616-22.
47. Kalia HS, Gaglio PJ. The Prevalence and Pathobiology of Nonalcoholic Fatty Liver Disease in Patients of Different Races or Ethnicities. Clinics in liver disease. 2016;20(2):215-24.

48. Schwimmer JB, Behling C, Newbury R, Deutsch R, Nievergelt C, Schork NJ, et al. Histopathology of pediatric nonalcoholic fatty liver disease. Hepatology. 2005;42(3):641-9.
49. Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. Pediatrics. 2006;118(4):1388-93.
50. Angulo P. Nonalcoholic fatty liver disease and liver transplantation. Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society. 2006;12(4):523-34.
51. Argo CK, Caldwell SH. Epidemiology and natural history of non-alcoholic steatohepatitis. Clinics in liver disease. 2009;13(4):511-31.
52. Wong RJ, Aguilar M, Cheung R, Perumpail RB, Harrison SA, Younossi ZM, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology. 2015;148(3):547-55.
53. Argo CK, Northup PG, Al-Osaimi AM, Caldwell SH. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. Journal of hepatology. 2009;51(2):371-9.
54. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2015;149(2):389-97 e10.
55. Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology. 2015;61(5):1547-54.
56. Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology. 2006;44(4):865-73.
57. Stepanova M, Rafiq N, Makhlouf H, Agrawal R, Kaur I, Younoszai Z, et al. Predictors of all-cause mortality and liver-related mortality in patients with non-alcoholic fatty liver disease (NAFLD). Digestive diseases and sciences. 2013;58(10):3017-23.
58. Ratziu V, Giral P, Charlotte F, Bruckert E, Thibault V, Theodorou I, et al. Liver fibrosis in overweight patients. Gastroenterology. 2000;118(6):1117-23.
59. Jepsen P, Vilstrup H, Mellemkjaer L, Thulstrup AM, Olsen JH, Baron JA, et al. Prognosis of patients with a diagnosis of fatty liver--a registry-based cohort study. Hepatogastroenterology. 2003;50(54):2101-4.
60. Dam-Larsen S, Franzmann M, Andersen IB, Christoffersen P, Jensen LB, Sorensen TI, et al. Long term prognosis of fatty liver: risk of chronic liver disease and death. Gut. 2004;53(5):750-5.
61. Angulo P, Keach JC, Batts KP, Lindor KD. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. Hepatology. 1999;30(6):1356-62.
62. Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2015;13(4):643-54 e1-9; quiz e39-40.
63. Marchesini G, Bugianesi E, Forlani G, Cerrelli F, Lenzi M, Manini R, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. Hepatology. 2003;37(4):917-23.

64. Bedogni G, Miglioli L, Masutti F, Castiglione A, Croce LS, Tiribelli C, et al. Incidence and natural course of fatty liver in the general population: the Dionysos study. Hepatology. 2007;46(5):1387-91.
65. Jakobsen MU, Berentzen T, Sorensen TI, Overvad K. Abdominal obesity and fatty liver. Epidemiol Rev. 2007;29:77-87.
66. Tolman KG, Fonseca V, Tan MH, Dalpiaz A. Narrative review: hepatobiliary disease in type 2 diabetes mellitus. Annals of internal medicine. 2004;141(12):946-56.
67. Kotronen A, Juurinen L, Hakkarainen A, Westerbacka J, Corner A, Bergholm R, et al. Liver fat is increased in type 2 diabetic patients and underestimated by serum alanine aminotransferase compared with equally obese nondiabetic subjects. Diabetes care. 2008;31(1):165-9.
68. Bhala N, Angulo P, van der Poorten D, Lee E, Hui JM, Saracco G, et al. The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study. Hepatology. 2011;54(4):1208-16.
69. Hui JM, Kench JG, Chitturi S, Sud A, Farrell GC, Byth K, et al. Long-term outcomes of cirrhosis in nonalcoholic steatohepatitis compared with hepatitis C. Hepatology. 2003;38(2):420-7.
70. Sanyal AJ, Banas C, Sargeant C, Luketic VA, Sterling RK, Stravitz RT, et al. Similarities and differences in outcomes of cirrhosis due to nonalcoholic steatohepatitis and hepatitis C. Hepatology. 2006;43(4):682-9.
71. Yatsuji S, Hashimoto E, Tobari M, Taniai M, Tokushige K, Shiratori K. Clinical features and outcomes of cirrhosis due to non-alcoholic steatohepatitis compared with cirrhosis caused by chronic hepatitis C. Journal of gastroenterology and hepatology. 2009;24(2):248-54.
72. Paradis V, Zalinski S, Chelbi E, Guedj N, Degos F, Vilgrain V, et al. Hepatocellular carcinomas in patients with metabolic syndrome often develop without significant liver fibrosis: a pathological analysis. Hepatology. 2009;49(3):851-9.
73. Fraser A, Harris R, Sattar N, Ebrahim S, Davey Smith G, Lawlor DA. Alanine aminotransferase, gamma-glutamyltransferase, and incident diabetes: the British Women's Heart and Health Study and meta-analysis. Diabetes care. 2009;32(4):741-50.
74. Targher G, Bertolini L, Padovani R, Rodella S, Tessari R, Zenari L, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes care. 2007;30(5):1212-8.
75. Targher G, Marra F, Marchesini G. Increased risk of cardiovascular disease in non-alcoholic fatty liver disease: causal effect or epiphenomenon? Diabetologia. 2008;51(11):1947-53.
76. Fracanzani AL, Burdick L, Raselli S, Pedotti P, Grigore L, Santorelli G, et al. Carotid artery intima-media thickness in nonalcoholic fatty liver disease. The American journal of medicine. 2008;121(1):72-8.
77. Fracanzani AL, Tiraboschi S, Pisano G, Consonni D, Baragetti A, Bertelli C, et al. Progression of carotid vascular damage and cardiovascular events in non-alcoholic fatty liver disease patients compared to the general population during 10 years of follow-up. Atherosclerosis. 2016;246:208-13.
78.

Nascimbeni F, Pais R, Bellentani S, Day CP, Ratziu V, Loria P, et al. From NAFLD in clinical practice to answers from guidelines. Journal of hepatology. 2013.

79. Ruffillo G, Fassio E, Alvarez E, Landeira G, Longo C, Dominguez N, et al. Comparison of NAFLD fibrosis score and BARD score in predicting fibrosis in nonalcoholic fatty liver disease. Journal of hepatology. 2011;54(1):160-3.
80. Feldstein AE, Wieckowska A, Lopez AR, Liu YC, Zein NN, McCullough AJ. Cytokeratin-18 fragment levels as noninvasive biomarkers for nonalcoholic steatohepatitis: a multicenter validation study. Hepatology. 2009;50(4):1072-8.
81. Kinner S, Reeder SB, Yokoo T. Quantitative Imaging Biomarkers of NAFLD. Digestive diseases and sciences. 2016;61(5):1337-47.
82. Tobari M, Hashimoto E, Yatsuji S, Torii N, Shiratori K. Imaging of nonalcoholic steatohepatitis: advantages and pitfalls of ultrasonography and computed tomography. Internal medicine. 2009;48(10):739-46.
83. Dasarathy S, Dasarathy J, Khiyami A, Joseph R, Lopez R, McCullough AJ. Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study. Journal of hepatology. 2009;51(6):1061-7.
84. Cengiz M, Senturk S, Cetin B, Bayrak AH, Bilek SU. Sonographic assessment of fatty liver: intraobserver and interobserver variability. International journal of clinical and experimental medicine. 2014;7(12):5453-60.
85. Afdhal NH, Bacon BR, Patel K, Lawitz EJ, Gordon SC, Nelson DR, et al. Accuracy of fibroscan, compared with histology, in analysis of liver fibrosis in patients with hepatitis B or C: a United States multicenter study. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2015;13(4):772-9 e1-3.
86. Friedrich-Rust M, Ong MF, Martens S, Sarrazin C, Bojunga J, Zeuzem S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology. 2008;134(4):960-74.
87. Murtagh J, Foerster V. Transient elastography (FibroScan) for non-invasive assessment of liver fibrosis. Issues in emerging health technologies. 2006(90):1-4.
88. Kwok R, Tse YK, Wong GL, Ha Y, Lee AU, Ngu MC, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease--the role of transient elastography and plasma cytokeratin-18 fragments. Alimentary pharmacology & therapeutics. 2014;39(3):254-69.
89. Castera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology. 2010;51(3):828-35.
90. Cassinotto C, Boursier J, de Ledinghen V, Lebigot J, Lapuyade B, Cales P, et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. Hepatology. 2016;63(6):1817-27.
91. Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. The American journal of gastroenterology. 2012;107(12):1862-71.

92. Hannah WN, Jr., Harrison SA. Noninvasive imaging methods to determine severity of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Hepatology. 2016;64(6):2234-43.
93. Fierbinteanu Braticevici C, Sporea I, Panaitescu E, Tribus L. Value of Acoustic Radiation Force Impulse Imaging Elastography for Non-invasive Evaluation of Patients with Nonalcoholic Fatty Liver Disease. Ultrasound in medicine & biology. 2013.
94. de Ledinghen V, Vergniol J, Foucher J, El-Hajbi F, Merrouche W, Rigalleau V. Feasibility of liver transient elastography with FibroScan using a new probe for obese patients. Liver international : official journal of the International Association for the Study of the Liver. 2010;30(7):1043-8.
95. Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology. 2012;55(1):199-208.
96. Singh S, Venkatesh SK, Loomba R, Wang Z, Sirlin C, Chen J, et al. Magnetic resonance elastography for staging liver fibrosis in non-alcoholic fatty liver disease: a diagnostic accuracy systematic review and individual participant data pooled analysis. European radiology. 2016;26(5):1431-40.
97. Loomba R, Wolfson T, Ang B, Hooker J, Behling C, Peterson M, et al. Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: a prospective study. Hepatology. 2014;60(6):1920-8.
98. Wang QB, Zhu H, Liu HL, Zhang B. Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: A meta-analysis. Hepatology. 2012;56(1):239-47.
99. Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, et al. Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography. Gastroenterology. 2016;150(3):626-37 e7.
100. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver D. Liver biopsy. Hepatology. 2009;49(3):1017-44.
101. Bravo AA, Sheth SG, Chopra S. Liver biopsy. The New England journal of medicine. 2001;344(7):495-500.
102. Myers RP, Fong A, Shaheen AA. Utilization rates, complications and costs of percutaneous liver biopsy: a population-based study including 4275 biopsies. Liver international : official j ournal of the International Association for the Study of the Liver. 2008;28(5):705-12.
103. Merriman RB, Ferrell LD, Patti MG, Weston SR, Pabst MS, Aouizerat BE, et al. Correlation of paired liver biopsies in morbidly obese patients with suspected nonalcoholic fatty liver disease. Hepatology. 2006;44(4):874-80.
104. Promrat K, Kleiner DE, Niemeier HM, Jackvony E, Kearns M, Wands JR, et al. Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. Hepatology. 2010;51(1):121-9.
105. Shojaee-Moradie F, Baynes KC, Pentecost C, Bell JD, Thomas EL, Jackson NC, et al. Exercise training reduces fatty acid availability and improves the insulin sensitivity of glucose metabolism. Diabetologia. 2007;50(2):404-13.
106. Johnson NA, Sachinwalla T, Walton DW, Smith K, Armstrong A, Thompson MW, et al. Aerobic exercise training reduces hepatic and visceral lipids in obese individuals without weight loss. Hepatology. 2009;50(4):1105-12.

107. van der Heijden GJ, Wang ZJ, Chu ZD, Sauer PJ, Haymond MW, Rodriguez LM, et al. A 12-week aerobic exercise program reduces hepatic fat accumulation and insulin resistance in obese, Hispanic adolescents. Obesity. 2010;18(2):384-90.
108. Marchesini G, Brizi M, Bianchi G, Tomassetti S, Zoli M, Melchionda N. Metformin in non-alcoholic steatohepatitis. Lancet. 2001;358(9285):893-4.
109. Nair S, Diehl AM, Wiseman M, Farr GH, Jr., Perrillo RP. Metformin in the treatment of non-alcoholic steatohepatitis: a pilot open label trial. Alimentary pharmacology & therapeutics. 2004;20(1):23-8.
110. Shields WW, Thompson KE, Grice GA, Harrison SA, Coyle WJ. The Effect of Metformin and Standard Therapy versus Standard Therapy alone in Nondiabetic Patients with Insulin Resistance and Nonalcoholic Steatohepatitis (NASH): A Pilot Trial. Therapeutic advances in gastroenterology. 2009;2(3):157-63.
111. Omer Z, Cetinkalp S, Akyildiz M, Yilmaz F, Batur Y, Yilmaz C, et al. Efficacy of insulin-sensitizing agents in nonalcoholic fatty liver disease. European journal of gastroenterology & hepatology. 2010;22(1):18-23.
112. Haukeland JW, Konopski Z, Eggesbo HB, von Volkmann HL, Raschpichler G, Bjoro K, et al. Metformin in patients with non-alcoholic fatty liver disease: a randomized, controlled trial. Scandinavian journal of gastroenterology. 2009;44(7):853-60.
113. Belfort R, Harrison SA, Brown K, Darland C, Finch J, Hardies J, et al. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. The New England journal of medicine. 2006;355(22):2297-307.
114. Aithal GP, Thomas JA, Kaye PV, Lawson A, Ryder SD, Spendlove I, et al. Randomized, placebo-controlled trial of pioglitazone in nondiabetic subjects with nonalcoholic steatohepatitis. Gastroenterology. 2008;135(4):1176-84.
115. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. The New England journal of medicine. 2010;362(18):1675-85.
116. Lincoff AM, Wolski K, Nicholls SJ, Nissen SE. Pioglitazone and risk of cardiovascular events in patients with type 2 diabetes mellitus: a meta-analysis of randomized trials. JAMA : the journal of the American Medical Association. 2007;298(10):1180-8.
117. Lavine JE, Schwimmer JB, Van Natta ML, Molleston JP, Murray KF, Rosenthal P, et al. Effect of vitamin E or metformin for treatment of nonalcoholic fatty liver disease in children and adolescents: the TONIC randomized controlled trial. JAMA : the journal of the American Medical Association. 2011;305(16):1659-68.
118. Miller ER, 3rd, Pastor-Barriuso R, Dalal D, Riemersma RA, Appel LJ, Guallar E. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. Annals of internal medicine. 2005;142(1):37-46.
119. Klein EA, Thompson IM, Jr., Tangen CM, Crowley JJ, Lucia MS, Goodman PJ, et al. Vitamin E and the risk of prostate cancer: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). JAMA : the journal of the American Medical Association. 2011;306(14):1549-56.
120. Gerss J, Kopcke W. The questionable association of vitamin E supplementation and mortality--inconsistent results of different meta-analytic approaches. Cellular and molecular biology. 2009;55 Suppl:OL1111-20.

121. Lindor KD, Kowdley KV, Heathcote EJ, Harrison ME, Jorgensen R, Angulo P, et al. Ursodeoxycholic acid for treatment of nonalcoholic steatohepatitis: results of a randomized trial. Hepatology. 2004;39(3):770-8.
122. Tsutsui M, Tanaka N, Kawakubo M, Sheena Y, Horiuchi A, Komatsu M, et al. Serum fragmented cytokeratin 18 levels reflect the histologic activity score of nonalcoholic fatty liver disease more accurately than serum alanine aminotransferase levels. Journal of clinical gastroenterology. 2010;44(6):440-7.
123. Rinella ME, Loomba R, Caldwell SH, Kowdley K, Charlton M, Tetri B, et al. Controversies in the Diagnosis and Management of NAFLD and NASH. Gastroenterology & hepatology. 2014;10(4):219-27.
124. Musso G, Cassader M, Rosina F, Gambino R. Impact of current treatments on liver disease, glucose metabolism and cardiovascular risk in non-alcoholic fatty liver disease (NAFLD): a systematic review and meta-analysis of randomised trials. Diabetologia. 2012;55(4):885-904.
125. Musso G, Gambino R, Cassader M, Pagano G. A meta-analysis of randomized trials for the treatment of nonalcoholic fatty liver disease. Hepatology. 2010;52(1):79-104.
126. Farrell GC, Chitturi S, Lau GK, Sollano JD, Asia-Pacific Working Party on N. Guidelines for the assessment and management of non-alcoholic fatty liver disease in the Asia-Pacific region: executive summary. Journal of gastroenterology and hepatology. 2007;22(6):775-7.
127. Commission on Chronic Illness (1957) Chronic illness in the United States: Volume I. Prevention of chronic illness, Cambridge, Mass., Harvard University Press, p. 45
128. Cole P, Morrison AS. Basic issues in population screening for cancer. Journal of the National Cancer Institute. 1980;64(5):1263-72.
129. Bellentani S, Scaglioni F, Marino M, Bedogni G. Epidemiology of non-alcoholic fatty liver disease. Digestive diseases. 2010;28(1):155-61.
130. Black WC, Welch HG. Screening for disease. AJR American journal of roentgenology. 1997;168(1):3-11.
131. Obuchowski NA, Graham RJ, Baker ME, Powell KA. Ten criteria for effective screening: their application to multislice CT screening for pulmonary and colorectal cancers. AJR American journal of roentgenology. 2001;176(6):1357-62.
132. Ascha MS, Hanouneh IA, Lopez R, Tamimi TA, Feldstein AF, Zein NN. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology. 2010;51(6):1972-8.
133. Soderberg C, Stal P, Askling J, Glaumann H, Lindberg G, Marmur J, et al. Decreased survival of subjects with elevated liver function tests during a 28-year follow-up. Hepatology. 2010;51(2):595-602.
134. Dam-Larsen S, Becker U, Franzmann MB, Larsen K, Christoffersen P, Bendtsen F. Final results of a long-term, clinical follow-up in fatty liver patients. Scandinavian journal of gastroenterology. 2009;44(10):1236-43.
135. Machado MV, Cortez-Pinto H. Non-invasive diagnosis of non-alcoholic fatty liver disease. A critical appraisal. Journal of hepatology. 2013;58(5):1007-19.
136. Canadian Institute for Health Information. National Health Expenditure Trends, 1975 to 2016. (Ottawa, Ont: CIHI; 2016).

137. Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada [4th Edition]. Ottawa, 2017:30.
138. U.S. National Library of Medicine [Internet]. Maryland: Health Services Research & Public Health; 2014. [updated 2017 March 07; cited 2017 Feb 18]. Available from: https://www.nlm.nih.gov/nichsr/hta101/ta10107.html.
139. Institut national d'excellence en santé et en services sociaux. The drug evaluation process [Analysis in brief on the Internet]. Québec: INESSS; 2013 [cited 2017 July 10]. 2 p. Available from: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Inscription_medicaments/Processus/Evaluation-medic_EN_28082013.pdf. Accessed July 15, 2017.
140. Bravo Vergel Y, Sculpher M. Quality-adjusted life years. Practical neurology. 2008;8(3):175-82.
141. Ratcliffe J, Longworth L, Young T, Bryan S, Burroughs A, Buxton M, et al. Assessing health-related quality of life pre- and post-liver transplantation: a prospective multicenter study. Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society. 2002;8(3):263-70.
142. Nichol MB, Sengupta N, Globe DR. Evaluating quality-adjusted life years: estimation of the health utility index (HUI2) from the SF-36. Medical decision making : an international journal of the Society for Medical Decision Making. 2001;21(2):105-12.
143. Younossi ZM, Boparai N, McCormick M, Price LL, Guyatt G. Assessment of utilities and health-related quality of life in patients with chronic liver disease. The American journal of gastroenterology. 2001;96(2):579-83.
144. Briggs AC, K.; Sculpher M. . Decision modelling for health economic evaluation. 1st ed. USA: Oxford University Press, Inc. ; 2006. 224 p.
145. Laupacis A, Feeny D, Detsky AS, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 1992;146(4):473-81.
146. Grosse SD. Assessing cost-effectiveness in healthcare: history of the $50,000 per QALY threshold. Expert review of pharmacoeconomics & outcomes research. 2008;8(2):165-78.
147. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the $50,000-per-QALY threshold. The New England journal of medicine. 2014;371(9):796-7.
148. Detsky AS NG, Krahn MD, Naimark D, Redelmeier DA Med Decis Making. 1997 Apr-Jun; 17(2):123-5.
149. Ratziu V, Bonyhay L, Di Martino V, Charlotte F, Cavallaro L, Sayegh-Tainturier MH, et al. Survival, liver failure, and hepatocellular carcinoma in obesity-related cryptogenic cirrhosis. Hepatology. 2002;35(6):1485-93.
150. Starley BQ, Calcagno CJ, Harrison SA. Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection. Hepatology. 2010;51(5):1820-32.
151. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology. 2012;55(6):2005-23.
152. Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-

analysis. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2007;5(10):1214-20.
153. Castera L, Pinzani M. Biopsy and non-invasive methods for the diagnosis of liver fibrosis: does it take two to tango? Gut. 2010;59(7):861-6.
154. Mahady SE, Webster AC, Walker S, Sanyal A, George J. The role of thiazolidinediones in non-alcoholic steatohepatitis - a systematic review and meta analysis. Journal of hepatology. 2011;55(6):1383-90.
155. Mahady SE, Wong G, Craig JC, George J. Pioglitazone and vitamin E for nonalcoholic steatohepatitis: a cost utility analysis. Hepatology. 2012;56(6):2172-9.
156. Hashimoto E, Yatsuji S, Tobari M, Taniai M, Torii N, Tokushige K, et al. Hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Journal of gastroenterology. 2009;44 Suppl 19:89-95.
157. Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. Bmj. 2011;342:d1766.
158. Friedrich-Rust M, Romen D, Vermehren J, Kriener S, Sadet D, Herrmann E, et al. Acoustic radiation force impulse-imaging and transient elastography for non-invasive assessment of liver fibrosis and steatosis in NAFLD. Eur J Radiol. 2011.
159. West J, Card TR. Reduced mortality rates following elective percutaneous liver biopsies. Gastroenterology. 2010;139(4):1230-7.
160. Chalasani NP, Sanyal AJ, Kowdley KV, Robuck PR, Hoofnagle J, Kleiner DE, et al. Pioglitazone versus vitamin E versus placebo for the treatment of non-diabetic patients with non-alcoholic steatohepatitis: PIVENS trial design. Contemporary clinical trials. 2009;30(1):88-96.
161. Miller DK, Homan SM. Determining transition probabilities: confusion and suggestions. Medical decision making : an international journal of the Society for Medical Decision Making. 1994;14(1):52-8.
162. Canadian Medical Association. Canadian Provincial Billing Guide. Health Canada, Government of Canada. Available at http://www.cma.ca/practicemanagement/medical-billing. Accessed July, 2013.
163. Steadman RL, L. Lorenzetti, D. Noseworthy, T. Rose, S. Sutherland, L. Clement, F. A Health Technology Assessment of Transient Elastography in Liver Disease. In: Unit UoCHTA, editor. Alberta: Alberta Health Technologies Decision Process; 2012.
164. Barry MH. Magnetic Resonance Imaging Needs Assessment In: Health NSDo, editor. Nova Scotia2004. p. 38.
165. Canadian Institute for Health Information. Medical Imaging in Canada, 2007 (Ottawa, Ont.: CIHI, 2008).
166. Peviva. M30 Apoptosense ELISA. Available at http://www.peviva.se/m30-apoptosense-elisa_us.aspx. Accessed July, 2013.
167. Canadian Institute for Health Information. The Cost of Acute Care Hospital Stays by Medical Condition in Canada, 2004-2005 (Ottawa: CIHI, 2008).
168. Thein HH, Isaranuwatchai W, Campitelli MA, Feld JJ, Yoshida E, Sherman M, et al. Health care costs associated with hepatocellular carcinoma: a population-based study. Hepatology. 2013;58(4):1375-84.
169. Taylor MC, Greig PD, Detsky AS, McLeod RS, Abdoh A, Krahn MD. Factors associated with the high cost of liver transplantation in adults. Canadian journal of surgery Journal canadien de chirurgie. 2002;45(6):425-34.

170. Bank of Canada Inflation Calculator. Available at http://www.bankofcanada.ca/rates/related/inflation-calculator. Accessed July, 2013.
171. David K, Kowdley KV, Unalp A, Kanwal F, Brunt EM, Schwimmer JB, et al. Quality of life in adults with nonalcoholic fatty liver disease: baseline data from the nonalcoholic steatohepatitis clinical research network. Hepatology. 2009;49(6):1904-12.
172. Chong CA, Gulamhussein A, Heathcote EJ, Lilly L, Sherman M, Naglie G, et al. Health-state utilities and quality of life in hepatitis C patients. The American journal of gastroenterology. 2003;98(3):630-8.
173. McLernon DJ, Dillon J, Donnan PT. Health-state utilities in liver disease: a systematic review. Medical decision making : an international journal of the Society for Medical Decision Making. 2008;28(4):582-92.
174. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. Bmj. 2012;344:e3502.
175. Yoneda M, Suzuki K, Kato S, Fujita K, Nozaki Y, Hosono K, et al. Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. Radiology. 2010;256(2):640-7.
176. Bota S, Herkner H, Sporea I, Salzl P, Sirli R, Neghina AM, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver international : official journal of the International Association for the Study of the Liver. 2013;33(8):1138-47.
177. Bambha K, Belt P, Abraham M, Wilson LA, Pabst M, Ferrell L, et al. Ethnicity and nonalcoholic fatty liver disease. Hepatology. 2012;55(3):769-80.
178. Chen J, Zhu Y, Zheng Q, Jiang J. Serum cytokeratin-18 in the diagnosis of non-alcoholic steatohepatitis: A meta-analysis. Hepatology research : the official journal of the Japan Society of Hepatology. 2013.
179. Chen J, Talwalkar JA, Yin M, Glaser KJ, Sanderson SO, Ehman RL. Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. Radiology. 2011;259(3):749-56.
180. Abner EL, Schmitt FA, Mendiondo MS, Marcum JL, Kryscio RJ. Vitamin E and all-cause mortality: a meta-analysis. Current aging science. 2011;4(2):158-70.
181. Mamtani R, Haynes K, Bilker WB, Vaughn DJ, Strom BL, Glanz K, et al. Association between longer therapy with thiazolidinediones and risk of bladder cancer: a cohort study. Journal of the National Cancer Institute. 2012;104(18):1411-21.
182. Vallarino C, Perez A, Fusco G, Liang H, Bron M, Manne S, et al. Comparing pioglitazone to insulin with respect to cancer, cardiovascular and bone fracture endpoints, using propensity score weights. Clinical drug investigation. 2013;33(9):621-31.
183. Monami M, Dicembrini I, Mannucci E. Thiazolidinediones and cancer: results of a meta-analysis of randomized clinical trials. Acta diabetologica. 2014;51(1):91-101.
184. Lazo M, Hernaez R, Eberhardt MS, Bonekamp S, Kamel I, Guallar E, et al. Prevalence of nonalcoholic fatty liver disease in the United States: the Third National Health and Nutrition Examination Survey, 1988-1994. American journal of epidemiology. 2013;178(1):38-45.
185. Angulo P. Nonalcoholic fatty liver disease. The New England journal of medicine. 2002;346(16):1221-31.
186. Preiss D, Sattar N. Non-alcoholic fatty liver disease: an overview of prevalence, diagnosis, pathogenesis and treatment considerations. Clinical science. 2008;115(5):141-50.

187. Wanless IR, Lentz JS. Fatty liver hepatitis (steatohepatitis) and obesity: an autopsy study with analysis of risk factors. Hepatology. 1990;12(5):1106-10.
188. Silverman JF, Pories WJ, Caro JF. Liver pathology in diabetes mellitus and morbid obesity. Clinical, pathological, and biochemical considerations. Pathology annual. 1989;24 Pt 1:275-302.
189. Pinto HC, Baptista A, Camilo ME, Valente A, Saragoca A, de Moura MC. Nonalcoholic steatohepatitis. Clinicopathological comparison with alcoholic hepatitis in ambulatory and hospitalized patients. Digestive diseases and sciences. 1996;41(1):172-9.
190. World Gastroenterology Organisation. WGO Global Guidelines Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. 2012:29.
191. Zelber-Sagi S, Lotan R, Shlomai A, Webb M, Harrari G, Buch A, et al. Predictors for incidence and remission of NAFLD in the general population during a seven-year prospective follow-up. Journal of hepatology. 2012;56(5):1145-51.
192. Teli MR, James OF, Burt AD, Bennett MK, Day CP. The natural history of nonalcoholic fatty liver: a follow-up study. Hepatology. 1995;22(6):1714-9.
193. Pais R, Pascale A, Fedchuck L, Charlotte F, Poynard T, Ratziu V. Progression from isolated steatosis to steatohepatitis and fibrosis in nonalcoholic fatty liver disease. Gastroenterologie clinique et biologique. 2010.
194. Wong VW, Wong GL, Choi PC, Chan AW, Li MK, Chan HY, et al. Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. Gut. 2010;59(7):969-74.
195. Rafiq N, Bai C, Fang Y, Srishord M, McCullough A, Gramlich T, et al. Long-term follow-up of patients with nonalcoholic fatty liver. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2009;7(2):234-8.
196. Angulo P. Long-term mortality in nonalcoholic fatty liver disease: is liver histology of any prognostic significance? Hepatology. 2010;51(2):373-5.
197. Fassio E, Alvarez E, Dominguez N, Landeira G, Longo C. Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies. Hepatology. 2004;40(4):820-6.
198. Hui AY, Wong VW, Chan HL, Liew CT, Chan JL, Chan FK, et al. Histological progression of non-alcoholic fatty liver disease in Chinese patients. Alimentary pharmacology & therapeutics. 2005;21(4):407-13.
199. McCullough AJ. The clinical features, diagnosis and natural history of nonalcoholic fatty liver disease. Clinics in liver disease. 2004;8(3):521-33, viii.
200. Marrero JA, Fontana RJ, Barrat A, Askari F, Conjeevaram HS, Su GL, et al. Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. Hepatology. 2005;41(4):707-16.
201. Altekruse SF, McGlynn KA, Reichman ME. Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(9):1485-91.
202. Greten TF, Papendorf F, Bleck JS, Kirchhoff T, Wohlberedt T, Kubicka S, et al. Survival rate in patients with hepatocellular carcinoma: a retrospective analysis of 389 patients. British journal of cancer. 2005;92(10):1862-8.

203. Lang H, Sotiropoulos GC, Brokalaki EI, Schmitz KJ, Bertona C, Meyer G, et al. Survival and recurrence rates after resection for hepatocellular carcinoma in noncirrhotic livers. Journal of the American College of Surgeons. 2007;205(1):27-36.
204. Perry JF, Charlton B, Koorey DJ, Waugh RC, Gallagher PJ, Crawford MD, et al. Outcome of patients with hepatocellular carcinoma referred to a tertiary centre with availability of multiple treatment options including cadaveric liver transplantation. Liver international : official journal of the International Association for the Study of the Liver. 2007;27(9):1240-8.
205. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology. 2007;45(4):846-54.
206. Friedrich-Rust M, Nierhoff J, Lupsor M, Sporea I, Fierbinteanu-Braticevici C, Strobel D, et al. Performance of Acoustic Radiation Force Impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. Journal of viral hepatitis. 2012;19(2):e212-9.
207. Nierhoff J, Chavez Ortiz AA, Herrmann E, Zeuzem S, Friedrich-Rust M. The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: a meta-analysis. European radiology. 2013;23(11):3040-53.
208. Carlson JJ, Kowdley KV, Sullivan SD, Ramsey SD, Veenstra DL. An evaluation of the potential cost-effectiveness of non-invasive testing strategies in the diagnosis of significant liver fibrosis. Journal of gastroenterology and hepatology. 2009;24(5):786-91.
209. DeKoven, M. Cost comparison: liver biopsy versus abdominal MRI. Memo to GE Healthcare from IMS Health Incorporated 2010.
210. Régie de l'assurance maladie du Québec (RAMQ). RAMQ liste de medicaments 2013. Quebec Ministry of Health and Social Services; 2013. Available at https://www.prod.ramq.gouv.qc.ca/DPI/PO/Commun/PDF/Liste_Med/Liste_Med/liste_med_cor_2013_07_15_fr.pdf. Accessed August, 2013.
211. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Jackson S, et al. Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. British journal of cancer. 2008;98(7):1166-75.
212. Aberg F, Maklin S, Rasanen P, Roine RP, Sintonen H, Koivusalo AM, et al. Cost of a quality-adjusted life year in liver transplantation: the influence of the indication and the model for end-stage liver disease score. Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society. 2011;17(11):1333-43.
213. Thoma C, Day CP, Trenell MI. Lifestyle interventions for the treatment of non-alcoholic fatty liver disease in adults: a systematic review. Journal of hepatology. 2012;56(1):255-66.

Appendix
Scientific Contribution

Cost-Utility Analysis of Nonalcoholic Steatohepatitis Screening

Submission Type: Original Article

Short title: Cost-utility Analysis of NASH Screening

Authors:

Eric Zhang, MD [1,5]

Claire Wartelle-Bladou, MD [2]

Luigi Lepanto, MD, MSc [1,3]

Jean Lachaine, PhD [4]

Guy Cloutier, PhD, PEng[1,5]

An Tang, MD, MSc [1,5]

Affiliations:

[1] Department of Radiology, Radio-Oncology and Nuclear Medicine, University of Montreal, Saint-Luc Hospital, Montreal, Quebec, Canada.

[2] Department of Gastroenterology and Hepatology, University of Montreal, Saint-Luc Hospital, Montreal, Quebec, Canada.

[3] Health Technology Assessment Unit, University of Montreal, Saint-Luc Hospital, Montreal, Quebec, Canada.

[4] Faculty of Pharmacy and Pharmacoeconomics, University of Montreal, Montreal, Quebec, Canada.

[5] Laboratory of Biorheology and Medical Ultrasonics, University of Montreal Hospital Research Center (CRCHUM), Montreal, Quebec, Canada.

Corresponding author information:

An Tang, MD MSc

Department of Radiology

Faculty of Medicine

University of Montreal, Saint-Luc Hospital

1058 rue Saint-Denis

Montréal, Québec Canada

H2X 3J4

Author Contributions

Eric Zhang: Analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, approved final submission.
Claire Wartelle-Bladou: Critical revision of the manuscript, approved final submission.
Luigi Lepanto: Study concept and design, critical revision of the manuscript, approved final submission.
Jean Lachaine: Critical revision of the manuscript, approved final submission.
Guy Cloutier: Critical revision of the manuscript, approved final submission.
An Tang: Study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, obtained funding, study supervision, approved final submission.

ABSTRACT

OBJECTIVES: Nonalcoholic fatty liver disease (NAFLD) is the most common liver disease in Western countries. No studies have examined the cost-effectiveness of screening for nonalcoholic steatohepatitis (NASH), its advanced form. **METHODS:** We performed a cost-utility analysis of annual non-invasive screening strategies using third-party payer perspective in a general population and compared it to screening in a high-risk obese or diabetic population. Screening algorithms involved well-studied techniques including NAFLD fibrosis score, transient elastography (TE), and acoustic radiation force impulse (ARFI) imaging for detecting advanced fibrosis (≥ F3); and plasma cytokeratin-18 for NASH detection. Liver biopsy and magnetic resonance elastography (MRE) were compared as confirmation methods. Canadian dollar costs were adjusted for inflation and discounted at 5%. Incremental cost-effectiveness ratio (ICER) of ≤$C50,000 was considered cost-effective. **RESULTS:** Compared with no screening, screening with NAFLD fibrosis score/TE/CK-18 algorithm with MRE as confirmation for advanced fibrosis had an ICER of $C26,143 per quality-adjusted life year (QALY) gained. Screening in high-risk obese or diabetic populations was more cost-effective, with an ICER of $C9,051 and $C7,991 per QALY gained respectively. Liver biopsy confirmation was not found to be cost-effective. **CONCLUSIONS:** Our model suggests that annual NASH screening in high-risk obese or diabetic populations can be cost-effective.

KEYWORDS: Cost-effectiveness; Nonalcoholic fatty liver disease (NAFLD); Fibrosis; Elastography; Screening

INTRODUCTION

Over the last decade, nonalcoholic fatty liver disease (NAFLD) has been recognized as the most prevalent liver disease in Western countries, due in large part to the high rates of obesity and type 2 diabetes [1]. It affects an estimated 20-30% of the general adult population, and as much as 90% of diabetic or obese patients [2-4]. The more advanced form, nonalcoholic steatohepatitis (NASH) may evolve to fibrosis, cirrhosis, liver failure, and hepatocellular carcinoma (HCC) [5-9].

Although healthcare costs related to NASH have not been well studied, they are estimated to be significant due to potential progression to liver failure and HCC. NASH-related liver failure is predicted to become the main cause of liver transplantation within the next decade [10]. Current practice guidelines do not advocate screening of NAFLD or NASH at this time, in part due to a lack of knowledge regarding optimal non-invasive diagnostic strategies, long-term benefits, and the cost-effectiveness of screening [11; 12]. Although liver biopsy is the current reference standard for diagnosis of steatohepatitis and advanced fibrosis in patients with NAFLD [12], its invasiveness makes it an unlikely modality for large-scale screening [13]. To address this issue, non-invasive blood tests and elastography methods have been introduced for the detection of NASH or advanced fibrosis ($\geq$ F3) [14-16]. Screening for this highly prevalent disease may be worthwhile [17], but as of yet, the value of screening strategies for NASH has not been studied. Given that healthcare systems are limited in their financial resources, competing screening strategies should be analyzed for their cost-effectiveness. Cost-utility studies, which incorporate the widely applicable quality-adjusted life year (QALY), a measure of health outcome, should guide policy makers in their decision to implement new screening strategies.

Weight loss is currently the recommended standard of care for NASH [12; 18]. In more advanced disease, vitamin E has been recommended as pharmacotherapy in non-diabetic patients with histologically-proven NASH [11; 12]. Pioglitazone is another suggested pharmacotherapy based upon latest randomized-control trials and meta-analyses [19; 20]. It should be noted that there are long-term safety concerns with vitamin E related to all-cause mortality and prostate cancer, and about thiazolidinediones related to bladder cancer, osteoporosis, and congestive heart failure [11; 12]. Despite these concerns, a recent cost-utility study established the cost-effectiveness of

pharmacological therapy for delaying the progression of NASH fibrosis using pioglitazone and vitamin E [21].

To our knowledge, there is currently no cost-utility study for NASH screening in the Western population. In this era of cost-containment, we believe it will be important to explore the cost-effectiveness and opportunity cost of screening for NASH. Thus, the primary aim of our study was to estimate the cost-effectiveness in a general population of different screening strategies for NASH or advanced fibrosis (≥ F3) detection, while incorporating currently recommended treatment practices. In order to determine the optimal population to target for screening, our secondary aim was to estimate the cost-effectiveness of these screening strategies in high-risk obese or type 2 diabetes populations.

MATERIALS AND METHODS

Markov Model and Assumptions

From a health-care system perspective, a decisional Markov model was developed (TreeAge Software, Williamstown, MA) to estimate the expected lifetime costs and QALYs associated with screening strategies for NASH. This model was constructed to mirror the natural history of NAFLD disease progression through the histopathological continuum of simple steatosis, NASH, fibrosis stages, and cirrhosis [5]. Patients with cirrhosis may progress onto liver failure and also have increased probability of developing HCC (**Figure 1**) [22].

To address our research aims, we ran the simulation for a general population and for high-risk populations, either with obesity or type 2 diabetes. Patients began screening at the age of 30. At the beginning of the simulation, the population was divided among these mutually exclusive health states according to mean prevalence rates reported in developed countries for a general population, an obese population, and a type 2 diabetes population, respectively. The model assumed an annual cycle length. In each cycle, simulated populations could remain in their health states or progress according to transition probabilities derived from literature. Screening and treatment strategies were superimposed onto this life cycle model of NAFLD. For the purpose of developing this model, histological improvement was assumed a good correlate for clinical outcomes. Both all-cause and liver-related mortalities were taken into account at each stage of disease. The simulation ended once every member of the population died. A lifetime horizon was chosen for this model to better reflect NAFLD disease progression [23], as well as to better represent the magnitude of costs and utilities associated with the disease. Peer-reviewed guidelines for economic evaluations were followed in the creation of this model [24; 25].

Competing Screening Strategies

The competing screening strategies incorporated independently and widely studied non-invasive tests. Plasma cytokeratin-18 (CK-18) was assessed for the noninvasive detection of NASH. NAFLD fibrosis score, ultrasound transient elastography (TE), and ultrasound acoustic radiation force impulse (ARFI) imaging were assessed for the detection of advanced fibrosis (≥ F3) [14; 16; 26]. We compared a sequential algorithm [14] that incorporates the NAFLD fibrosis score, transient elastography, and CK-18 with biopsy confirmation to no screening in our Markov model.

The underlying assumption was that a strategy combining noninvasive methods for NASH and fibrosis detection would decrease the number of unnecessary liver biopsies [14; 16]. Given the similar sensitivity of ultrasound-based elastography for the detection of advanced fibrosis, we compared a variant of this sequential algorithm by substituting TE with ARFI [27]. In addition, considering the high diagnostic accuracy of magnetic resonance elastography (MRE) for fibrosis staging [26], we also compared MRE against liver biopsy for the confirmation of advanced fibrosis. The mortality risk associated with liver biopsy as well as the costs associated with severe bleeding complications were implemented in the model [28]. **Figure 2** illustrates the various screening strategies compared in our study.

Treatment Arms

Three treatment branches were implemented in the model. In accordance with international guidelines, NASH patients with no or mild fibrosis ($F \leq 1$) were treated with lifestyle intervention and weight loss, whereas patients with advanced fibrosis ($F \geq 3$) were treated pharmacologically [29]. Lifestyle intervention aimed to achieve an overall weight reduction of 7-10% by combining regimented exercise, diet, and behavior adjustments. The treatment effect on NASH progression was calculated from a randomized controlled trial looking at the histological improvements of a lifestyle intervention program versus standard of care [18]. The pharmacotherapies considered in our model included vitamin E and pioglitazone [19; 30]. Treatment effects on fibrosis progression were estimated by applying the relative risk for histological improvement used in a previous cost-utility analysis [21]. Pharmacotherapies were stopped in the event of liver decompensation development, in accordance with assumptions made previously [21].

Model Parameter Estimates

Prevalence, annual transition probabilities, and mortality risk for the Markov model were derived from a systematic literature review (**Table 1**). Annual transition probabilities were calculated based on the approach outlined by Miller and Homan for converting rates over time [31]. Screening test sensitivities were obtained from meta-analyses. Liver biopsy, as the accepted reference standard, was assumed to have 100% accuracy.

Costs

Annual healthcare costs were derived from the Canadian Provincial Billing Guides [32]. Relevant costs include primary care follow-up, specialist consultation, and blood work panels to rule out alternative diagnoses of chronic liver disease. Screening tests were micro-costed from the Canadian Provincial Billing Guides, the Canadian Agency of Drugs and Technologies in Health, and related literature on micro-costing of elastography methods in Canada [33-35]. Cost of cytokeratin-18 M30-apoptosense ELISA kit (PEVIVA, Bromma, Sweden) were obtained from the company website [36]. Annual patient care costs for liver decompensation were taken from the Canadian Institute of Health Information [37]. The costs of HCC management and liver transplantation were derived from published literature specific to the Canadian healthcare system [38; 39]. All costs incorporated into the model are in 2013 Canadian dollars ($CAD) (**Table 2 and Supplementary Table 1**). Costs were adjusted for inflation to 2013 when needed using the national inflation index [40].

Health-Related Quality of Life

For health-related quality of life data, we used the largest study performed in patients with NAFLD to date [41]. This study provided quality of life data on patients with NAFLD and NASH in the form of a SF-36 survey. The data from these surveys were then converted to utility estimates using the method described by Nichol et al. [42]. Further health-related quality of life information on NASH-associated fibrosis, cirrhosis and hepatic decompensation were nonexistent. Therefore, we used utilities from health-related quality of life studies on other causes of chronic liver disease [43-46]. Given the benign nature of the disease, simple steatosis was assumed an utility estimate of 1. Utility values for each health state are reported in **Table 3**.

Outcomes

Outcomes were measured in terms of costs ($CAD) and in terms of quality-adjusted life years gained (QALYs). The incremental cost-effectiveness ratio (ICER) of each strategy was calculated as the incremental difference in cost divided by the incremental difference in quality-adjusted life years of two consecutive strategies. In the Canadian heath care setting, ICERs of less than $C50,000 per QALY gained is usually considered cost-effective. The discount rate was set at 5% in accordance with Canadian guidelines [25]. A strategy is *dominating* when it results in lower

cost and higher QALYs in comparison to another and *dominated* when it results with higher cost and less QALYs in comparison to another.

Sensitivity Analyses

The robustness of our results was assessed in terms of one-way sensitivity analyses, in which all model parameters were varied across a range taken from published data or at 95% confidence intervals. For the transition probability from simple steatosis to NASH, which was not readily available due to a paucity of data, we took a large range of plausible values [23]. Two-way sensitivity analyses were performed on select pairs of parameters that were influential in one-way sensitivity analyses. There were not enough published data to build probability distributions for undergoing a probabilistic sensitivity analysis.

RESULTS

Table 4 illustrates the results for the top three dominating screening strategies for the base case analysis for each population studied. In the general population, no surveillance as a baseline strategy costs $C6,561 per person with a total utility value of 42.04 QALYs gained over the lifetime of the patient. NAFLD fibrosis score/TE/CK-18 sequential strategy with MRE confirmation for advanced fibrosis and vitamin E treatment cost $C3,136 more per person but also delivered incremental utility increase of 0.12 QALYs. This strategy was found to be cost-effective with an ICER of $C26,143/QALY gained according to a threshold of $C50,000/QALY gained. The same strategy with pioglitazone treatment was found to have an ICER of $C199,870/QALY gained.

Cost-utility in High-risk Populations

In an obese population, the NAFLD fibrosis score/TE/CK-18 sequential strategy with MRE confirmation for advanced fibrosis and vitamin E treatment resulted in an ICER of $C9,051/QALY gained compared to no surveillance. In a type 2 diabetic population, the same screening strategy resulted in an ICER of $C7,991/QALY gained compared to no surveillance. The remaining screening strategies not seen in **Table 4** were dominated and therefore not found to be cost-effective.

Sensitivity Analyses

One-way sensitivity analysis results for the NAFLD fibrosis score/TE/CK-18 sequential algorithm with MRE confirmation and vitamin E treatment are summarized in **Figure 3**. In this analysis, all parameters used during the simulations were varied through the range of values found in literature or by applying 95% confidence intervals to test the robustness of our results given the potential uncertainty of parameter values. The ICER for the base case scenario is delineated by the vertical line. The ICERs within the variable range tested move from the blue (lower range) to the red side (upper range).

Three variables were found to have the greatest effect on the ICER: the test cost for TE, the starting age for screening, and the annual transition probability of steatosis to NASH. If the cost of an

individual TE test was assumed to be $C50, then the ICER was $C20,521/QALY gained. At an upper limit assumption of $C250 per test, then the ICER increased to $C43,040/QALY gained. If the starting age for screening began at 18 years of age, the ICER was found to be as low as $C17,535/QALY gained. However, if screening began at 43 years old or later, the ICER surpassed the $C50,000/QALY gained threshold. The annual probability of developing steatohepatitis, a value that has not been well documented in literature, also had an important effect on the ICER. If the annual probability of developing steatohepatitis was 8.8%, then the ICER would be $C11,164/QALY gained; however, if the annual incidence of steatohepatitis were as low as 0.03%, then the ICER would increase to more than $C42,787/QALY gained.

In accordance with Canadian health technology assessment guidelines, the model was assessed using 0% as well as 3% discount rates for comparison purposes with other jurisdictions [25]. In general, lowering the discount rate resulted in more cost-effective strategies. At 0% discount, all screening strategies became more cost-effective. The NAFLD fibrosis score/TE/CK-18 sequential algorithm with MRE confirmation resulted in an ICER of $C15,493/QALY gained.

Two-way sensitivity analyses were conducted to examine the effect on ICER of varying pairs of influential variables simultaneously. This can help distinguish particular thresholds whereby one strategy becomes more cost-effective than another assuming a $C50,000/QALY gained threshold. Two-way sensitivity analyses found that ARFI and TE were interchangeable in the sequential algorithm.

DISCUSSION

We performed a cost-utility analysis to address the current knowledge gap regarding the cost-effectiveness of screening for NASH. Given the present inclination towards reducing over-diagnosis and over-treatment in the healthcare community [47], we believe it is important to estimate the cost-effectiveness of a NASH screening program when compared to the opportunity cost of no screening. Specifically, we analyzed screening strategies for the detection of steatohepatitis and NASH-fibrosis, two advanced forms of NAFLD that may progress to end-stage liver disease. At present, no single non-invasive test is accurate enough to replace liver biopsy which remains the established reference standard for the diagnosis of steatohepatitis and advanced fibrosis in patients with NAFLD [12]. However, due to the invasive and costly nature of liver biopsy, it is not feasible as a screening test. By combining the most widely studied non-invasive tests, we are able to categorize patients according to their probability of having advanced disease and thus limit the total number of liver biopsies [14; 16]. To further decrease the invasiveness of a screening strategy for NASH, we examined the potential of MRE as an alternate reference standard to liver biopsy for liver fibrosis diagnosis, based on promising meta-analysis results [26]. Finally, we compared these screening algorithms in both general and high-risk populations to determine the most cost-effective population to screen.

Our model suggests that, in a general population, a sequential algorithm that includes the NAFLD fibrosis score/TE/CK-18, with MRE confirmation for advanced fibrosis, and vitamin E as treatment, can be a cost-effective surveillance strategy with an ICER of $C26,143/QALY gained. In comparison, the same sequential algorithm with pioglitazone treatment was found to have a higher ICER of $C199,870/QALY gained. The results indicate that the combination of non-invasive tests for detection of advanced fibrosis and NASH, with lifestyle changes and vitamin E as treatment, provides incremental gains of QALYs over no surveillance. By detecting earlier stages of the NAFLD and by implementing treatment according to current guidelines, this surveillance strategy demonstrates the potential to limit the transition of patients towards liver cirrhosis and end-stage liver disease, and its associated quality-of-life and economic costs.

Vitamin E appears to be more cost-effective as a treatment than pioglitazone in our model. A previous cost-utility analysis by Mahady et al. suggested that pioglitazone was a more cost-effective treatment than vitamin E in NAFLD patients at risk of cirrhotic progression [21]. This difference may be reflective of the fact that vitamin E is significantly less costly than pioglitazone in Canada and that we also incorporated lifestyle treatment for NASH in our model. However, before we can reach a definitive conclusion, it is important to note that the absolute difference in costs and QALYs between the two treatment strategies is quite small. Given that there is less than a 1% difference in costs and a 0.01% difference in QALYs between the two strategies, we cannot conclude with absolute certainty that one strategy is significantly more cost-effective than another.

Our model suggests that MRE is more cost-effective than liver biopsy as a confirmation method in a screening program for advanced fibrosis (≥ F3). Strategies with liver biopsy as confirmation for advanced NASH-fibrosis were more costly for less QALYs gained. This result reflects both the potential of MRE as an alternative reference standard, as well as the mortality and morbidity associated with liver biopsy. In recent years, MRE has emerged as a highly accurate modality for the staging of liver fibrosis, with histopathology as the reference standard [26]. From the point of view of a screening program, confirmation with MRE would likely be better accepted by the general population, given that it is non-invasive. Liver biopsy is associated with a small, but significant, risk of mortality, as well as a morbidity risk associated with severe bleeding, which may limit the willingness of asymptomatic patients to undergo a screening program, as well as the number of prescriptions by physicians [48].

The one-way sensitivity analysis identified the key drivers of cost-effectiveness. These include the cost of TE, starting age of surveillance, and the annual transitional probability from simple steatosis to NASH. In the base case scenario, the underlying assumption was that screening would begin at 30 years of age. Given that NASH and its complications are becoming an increasing problem among younger people [12], earlier screening could be a possibility. Ultrasound-based elastography methods, namely TE or ARFI, have similar sensitivities for detection of advanced fibrosis [27; 49; 50] and may be used interchangeably for fibrosis staging in the clinical workflow [16]. However, in our model, a sequential algorithm in which TE was substituted with ARFI was found to be dominated by the leading screening strategy with TE. This difference in cost/QALY

gained may be explained by the higher cost of ARFI over TE in our micro-costing scenario. Since itemized costs for these elastography tests have yet to be established in the Canadian healthcare billing guides, the micro-costing relied on a series of assumptions. To address the inherent uncertainties surrounding our assumptions on costs for these exams, we performed a two-way sensitivity analysis, which suggested that ARFI and TE were close to equivalent in the sequential screening algorithm along the range of costs from $C50 to $C250.

Our secondary aim was to examine the cost-effectiveness of these same screening strategies in high-risk obese and type 2 diabetes populations. We found that the most cost-effective screening strategy in a general population ($C26,143/QALY gained) was significantly more cost-effective in high-risk populations ($C9,051/QALY gained in an obese population and $C7,991/QALY gained in a type 2 diabetes population). While we did not investigate specific higher-risk ethnic groups [51] in our current model, we suspect that screening in higher-risk ethnic populations would similarly be more cost-effective.

One of the principles of preventing over-diagnosis is to better differentiate between benign disease and progressive disease that will cause more harm [47]. Thus, in our model, we did not screen for simple steatosis because, without inflammation, it is considered a benign, non-progressive disease in the majority of patients and not likely to develop into advanced fibrosis during their lifetimes [52]. Instead, we focused on the non-invasive detection of steatohepatitis and fibrosis, both of which are progressive stages of NAFLD and can lead to major complications if not found and treated. In our model, the non-invasive detection of NASH without advanced fibrosis depended on CK-18 fragments, which has a fair accuracy for NASH screening [53], with confirmation by liver biopsy. The current challenge with CK-18 includes its limited availability and as such it has not been introduced in clinical practice in Canada. Alternatively, MRE has been proposed for detection of NASH [54]. However, this will require independent validation in the future before we can consider an entirely non-invasive screening algorithm.

There are limitations to our study. The relevance of screening relies on the assumption that effective long-term therapy for NASH exists. It is conceivable that the histological improvements observed in short-term randomized controlled trials on lifestyle modification [18], pioglitazone,

and vitamin E [19; 20] may not be sustainable after discontinuation of therapy and over the lifetime horizon. Thus, longer-term studies on NASH and antifibrotic treatment are required. Nonetheless, current guidelines suggest the usage of pharmacologic therapy (vitamin E and glitazones) with caution in specific patients with elevated risk of progression to cirrhosis who have failed lifestyle intervention [11].

Further, we did not model the potential side effects of pharmacotherapy. Glitazones have been implicated in long-term safety concerns regarding cardiovascular disease, bladder cancer, and bone loss whereas vitamin E has been associated with a possible increase in all-cause mortality and risk of prostate cancer. However, given that there has been much controversy and conflicting results in the literature [55-59], and that it was not possible to model all complications for the purposes of an economic model, we decided not to implement them. Thus, the QALYs gained per strategy may be less than reported in our results, and the ICERs of surveillance strategies may be less favorable.

Our study has the following strengths. The algorithms studied in our model were derived from meta-analyses and compatible with current guidelines. The model parameters were based on a systematic literature review to identify prevalence, transition probabilities, costs, and utilities. These parameters represent a comprehensive simulation of NAFLD continuum. Where possible, we used utility estimates for steatohepatitis derived from a population with NASH [41]. Using present literature, we were able to address questions surrounding the long-term cost-effectiveness of NASH screening in different at-risk groups and assess the driving factors of cost-effectiveness in our sensitivity analyses.

Although economic modeling can be very useful, it is important to highlight deficiencies in current data that would otherwise have enriched our model. Through our extensive literature review in building our model, we have identified gaps in knowledge that should be addressed in future studies. There is currently a lack of health-related quality of life studies in NASH and NASH-fibrosis. No definitive utility estimates are available for different disease states of NASH and estimates were taken from mathematical conversions of SF-36 surveys or from other causes of chronic liver disease. Long-term transition rates for specific high-risk populations or ethnicities were sparse and further epidemiological data are needed. Finally, long-term treatment for NASH

and NASH-fibrosis should be more definitely studied in multicenter randomized controlled trials, in particular regarding their long-term efficacy and side effects.

In summary, our cost-utility model suggests that NASH screening is cost-effective with non-invasive screening methods for steatohepatitis and advanced fibrosis. Furthermore, screening in high-risk populations of obese or type 2 diabetes patients is more cost-effective than in a general Western population. Before decision-makers decide to implement a screening program, further studies should better establish the quality of life in NASH and the long-term effectiveness and safety of therapy.

ACKNOWLEDGEMENTS

This work was supported by the (1) Canadian Institutes of Health Research - Institute of Nutrition, Metabolism and Diabetes (CIHR-IMND) Operating Grant (#301520), (2) Fonds de recherche du Québec en Santé (Junior 1 Career Award FRQS-ARQ #26993 to An Tang, MD), (3) Recruitment fund for new scientist at the CRCHUM to An Tang. Funding agencies did not have any role in the design and conduct of the study, collection, management, analysis or interpretation of the data, preparation, review, or approval of the manuscript. There is no conflict of interest.

REFERENCES

1 Neuschwander-Tetri BA, Caldwell SH (2003) Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. Hepatology 37:1202-1219

2 Bellentani S, Scaglioni F, Marino M, Bedogni G (2010) Epidemiology of non-alcoholic fatty liver disease. Dig Dis 28:155-161

3 Adams LA, Sanderson S, Lindor KD, Angulo P (2005) The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. J Hepatol 42:132-138

4 Browning JD, Szczepaniak LS, Dobbins R et al (2004) Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatology 40:1387-1395

5 Farrell GC, Larter CZ (2006) Nonalcoholic fatty liver disease: from steatosis to cirrhosis. Hepatology 43:S99-S112

6 Ekstedt M, Franzen LE, Mathiesen UL et al (2006) Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology 44:865-873

7 Ascha MS, Hanouneh IA, Lopez R, Tamimi TA, Feldstein AF, Zein NN (2010) The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. Hepatology 51:1972-1978

8 Ratziu V, Bonyhay L, Di Martino V et al (2002) Survival, liver failure, and hepatocellular carcinoma in obesity-related cryptogenic cirrhosis. Hepatology 35:1485-1493

9 Starley BQ, Calcagno CJ, Harrison SA (2010) Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection. Hepatology 51:1820-1832

10 Charlton MR, Burns JM, Pedersen RA, Watt KD, Heimbach JK, Dierkhising RA (2011) Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. Gastroenterology 141:1249-1253

11 Ratziu V, Bellentani S, Cortez-Pinto H, Day C, Marchesini G (2010) A position statement on NAFLD/NASH based on the EASL 2009 special conference. J Hepatol 53:372-384

12 Chalasani N, Younossi Z, Lavine JE et al (2012) The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology 55:2005-2023

13 Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver D (2009) Liver biopsy. Hepatology 49:1017-1044

14 Musso G, Gambino R, Cassader M, Pagano G (2011) Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. Ann Med 43:617-649

15 Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM (2007) Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-analysis. Clin Gastroenterol Hepatol 5:1214-1220

16 Machado MV, Cortez-Pinto H (2013) Non-invasive diagnosis of non-alcoholic fatty liver disease. A critical appraisal. J Hepatol 58:1007-1019

17 Castera L, Pinzani M (2010) Biopsy and non-invasive methods for the diagnosis of liver fibrosis: does it take two to tango? Gut 59:861-866

18 Promrat K, Kleiner DE, Niemeier HM et al (2010) Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. Hepatology 51:121-129

19 Sanyal AJ, Chalasani N, Kowdley KV et al (2010) Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. N Engl J Med 362:1675-1685
20 Mahady SE, Webster AC, Walker S, Sanyal A, George J (2011) The role of thiazolidinediones in non-alcoholic steatohepatitis - a systematic review and meta analysis. J Hepatol 55:1383-1390
21 Mahady SE, Wong G, Craig JC, George J (2012) Pioglitazone and vitamin E for nonalcoholic steatohepatitis: a cost utility analysis. Hepatology 56:2172-2179
22 Hashimoto E, Yatsuji S, Tobari M et al (2009) Hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. J Gastroenterol 44 Suppl 19:89-95
23 Vernon G, Baranova A, Younossi ZM (2011) Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 34:274-285
24 Petrou S, Gray A (2011) Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. BMJ 342:d1766
25 Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada [3rd Edition]. Ottawa, 2006:75.
26 Wang QB, Zhu H, Liu HL, Zhang B (2012) Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: A meta-analysis. Hepatology 56:239-247
27 Friedrich-Rust M, Romen D, Vermehren J et al (2011) Acoustic radiation force impulse-imaging and transient elastography for non-invasive assessment of liver fibrosis and steatosis in NAFLD. Eur J Radiol. 10.1016/j.ejrad.2011.10.029
28 West J, Card TR (2010) Reduced mortality rates following elective percutaneous liver biopsies. Gastroenterology 139:1230-1237
29 Nascimbeni F, Pais R, Bellentani S et al (2013) From NAFLD in clinical practice to answers from guidelines. J Hepatol. 10.1016/j.jhep.2013.05.044
30 Chalasani NP, Sanyal AJ, Kowdley KV et al (2009) Pioglitazone versus vitamin E versus placebo for the treatment of non-diabetic patients with non-alcoholic steatohepatitis: PIVENS trial design. Contemp Clin Trials 30:88-96
31 Miller DK, Homan SM (1994) Determining transition probabilities: confusion and suggestions. Med Decis Making 14:52-58
32 Canadian Medical Association. Canadian Provincial Billing Guide. Health Canada, Government of Canada. Available at http://www.cma.ca/practicemanagement/medical-billing. Accessed July, 2013.
33 Steadman RL, L. Lorenzetti, D. Noseworthy, T. Rose, S. Sutherland, L. Clement, F. (2012) A Health Technology Assessment of Transient Elastography in Liver Disease. In: Unit UoCHTA, (ed). Alberta Health Technologies Decision Process, Alberta
34 Barry MH (2004) Magnetic Resonance Imaging Needs Assessment In: Health NSDo, (ed), Nova Scotia, pp 38
35 Canadian Institute for Health Information. Medical Imaging in Canada, 2007 (Ottawa, Ont.: CIHI, 2008).
36 Peviva. M30 Apoptosense ELISA. Available at http://www.peviva.se/m30-apoptosense-elisa_us.aspx. Accessed July, 2013.
37 Canadian Institute for Health Information. The Cost of Acute Care Hospital Stays by Medical Condition in Canada, 2004-2005 (Ottawa: CIHI, 2008).

38 Thein HH, Isaranuwatchai W, Campitelli MA et al (2013) Health care costs associated with hepatocellular carcinoma: a population-based study. Hepatology 58:1375-1384
39 Taylor MC, Greig PD, Detsky AS, McLeod RS, Abdoh A, Krahn MD (2002) Factors associated with the high cost of liver transplantation in adults. Can J Surg 45:425-434
40 Bank of Canada Inflation Calculator. Available at http://www.bankofcanada.ca/rates/related/inflation-calculator. Accessed July, 2013.
41 David K, Kowdley KV, Unalp A et al (2009) Quality of life in adults with nonalcoholic fatty liver disease: baseline data from the nonalcoholic steatohepatitis clinical research network. Hepatology 49:1904-1912
42 Nichol MB, Sengupta N, Globe DR (2001) Evaluating quality-adjusted life years: estimation of the health utility index (HUI2) from the SF-36. Med Decis Making 21:105-112
43 Chong CA, Gulamhussein A, Heathcote EJ et al (2003) Health-state utilities and quality of life in hepatitis C patients. Am J Gastroenterol 98:630-638
44 Younossi ZM, Boparai N, McCormick M, Price LL, Guyatt G (2001) Assessment of utilities and health-related quality of life in patients with chronic liver disease. Am J Gastroenterol 96:579-583
45 Ratcliffe J, Longworth L, Young T et al (2002) Assessing health-related quality of life pre- and post-liver transplantation: a prospective multicenter study. Liver Transpl 8:263-270
46 McLernon DJ, Dillon J, Donnan PT (2008) Health-state utilities in liver disease: a systematic review. Med Decis Making 28:582-592
47 Moynihan R, Doust J, Henry D (2012) Preventing overdiagnosis: how to stop harming the healthy. BMJ 344:e3502
48 Fernandez-Salazar L, Velayos B, Aller R, Lozano F, Garrote JA, Gonzalez JM (2011) Percutaneous liver biopsy: patients' point of view. Scand J Gastroenterol 46:727-731
49 Yoneda M, Suzuki K, Kato S et al (2010) Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. Radiology 256:640-647
50 Bota S, Herkner H, Sporea I et al (2013) Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver Int 33:1138-1147
51 Bambha K, Belt P, Abraham M et al (2012) Ethnicity and nonalcoholic fatty liver disease. Hepatology 55:769-780
52 Argo CK, Caldwell SH (2009) Epidemiology and natural history of non-alcoholic steatohepatitis. Clin Liver Dis 13:511-531
53 Chen J, Zhu Y, Zheng Q, Jiang J (2013) Serum cytokeratin-18 in the diagnosis of non-alcoholic steatohepatitis: A meta-analysis. Hepatol Res. 10.1111/hepr.12197
54 Chen J, Talwalkar JA, Yin M, Glaser KJ, Sanderson SO, Ehman RL (2011) Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. Radiology 259:749-756
55 Abner EL, Schmitt FA, Mendiondo MS, Marcum JL, Kryscio RJ (2011) Vitamin E and all-cause mortality: a meta-analysis. Curr Aging Sci 4:158-170
56 Miller ER, 3rd, Pastor-Barriuso R, Dalal D, Riemersma RA, Appel LJ, Guallar E (2005) Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. Ann Intern Med 142:37-46
57 Mamtani R, Haynes K, Bilker WB et al (2012) Association between longer therapy with thiazolidinediones and risk of bladder cancer: a cohort study. J Natl Cancer Inst 104:1411-1421

58 Vallarino C, Perez A, Fusco G et al (2013) Comparing pioglitazone to insulin with respect to cancer, cardiovascular and bone fracture endpoints, using propensity score weights. Clin Drug Investig 33:621-631
59 Monami M, Dicembrini I, Mannucci E (2014) Thiazolidinediones and cancer: results of a meta-analysis of randomized clinical trials. Acta Diabetol 51:91-101
60 Lazo M, Hernaez R, Eberhardt MS et al (2013) Prevalence of nonalcoholic fatty liver disease in the United States: the Third National Health and Nutrition Examination Survey, 1988-1994. Am J Epidemiol 178:38-45
61 Targher G, Bertolini L, Padovani R et al (2007) Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care 30:1212-1218
62 Angulo P (2002) Nonalcoholic fatty liver disease. N Engl J Med 346:1221-1231
63 Medina J, Fernandez-Salazar LI, Garcia-Buey L, Moreno-Otero R (2004) Approach to the pathogenesis and treatment of nonalcoholic steatohepatitis. Diabetes Care 27:2057-2066
64 Preiss D, Sattar N (2008) Non-alcoholic fatty liver disease: an overview of prevalence, diagnosis, pathogenesis and treatment considerations. Clin Sci (Lond) 115:141-150
65 Williams CD, Stengel J, Asike MI et al (2011) Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. Gastroenterology 140:124-131
66 Wanless IR, Lentz JS (1990) Fatty liver hepatitis (steatohepatitis) and obesity: an autopsy study with analysis of risk factors. Hepatology 12:1106-1110
67 Silverman JF, Pories WJ, Caro JF (1989) Liver pathology in diabetes mellitus and morbid obesity. Clinical, pathological, and biochemical considerations. Pathol Annu 24 Pt 1:275-302
68 Pinto HC, Baptista A, Camilo ME, Valente A, Saragoca A, de Moura MC (1996) Nonalcoholic steatohepatitis. Clinicopathological comparison with alcoholic hepatitis in ambulatory and hospitalized patients. Dig Dis Sci 41:172-179
69 World Gastroenterology Organisation. WGO Global Guidelines Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. 2012:29.
70 Zelber-Sagi S, Lotan R, Shlomai A et al (2012) Predictors for incidence and remission of NAFLD in the general population during a seven-year prospective follow-up. J Hepatol 56:1145-1151
71 Teli MR, James OF, Burt AD, Bennett MK, Day CP (1995) The natural history of nonalcoholic fatty liver: a follow-up study. Hepatology 22:1714-1719
72 Pais R, Pascale A, Fedchuck L, Charlotte F, Poynard T, Ratziu V (2010) Progression from isolated steatosis to steatohepatitis and fibrosis in nonalcoholic fatty liver disease. Gastroenterol Clin Biol. 10.1016/j.gcb.2010.06.004
73 Wong VW, Wong GL, Choi PC et al (2010) Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. Gut 59:969-974
74 Rafiq N, Bai C, Fang Y et al (2009) Long-term follow-up of patients with nonalcoholic fatty liver. Clin Gastroenterol Hepatol 7:234-238
75 Angulo P (2010) Long-term mortality in nonalcoholic fatty liver disease: is liver histology of any prognostic significance? Hepatology 51:373-375
76 Fassio E, Alvarez E, Dominguez N, Landeira G, Longo C (2004) Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies. Hepatology 40:820-826

77 Hui AY, Wong VW, Chan HL et al (2005) Histological progression of non-alcoholic fatty liver disease in Chinese patients. Aliment Pharmacol Ther 21:407-413
78 Bhala N, Angulo P, van der Poorten D et al (2011) The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study. Hepatology 54:1208-1216
79 Hui JM, Kench JG, Chitturi S et al (2003) Long-term outcomes of cirrhosis in nonalcoholic steatohepatitis compared with hepatitis C. Hepatology 38:420-427
80 Sanyal AJ, Banas C, Sargeant C et al (2006) Similarities and differences in outcomes of cirrhosis due to nonalcoholic steatohepatitis and hepatitis C. Hepatology 43:682-689
81 Soderberg C, Stal P, Askling J et al (2010) Decreased survival of subjects with elevated liver function tests during a 28-year follow-up. Hepatology 51:595-602
82 McCullough AJ (2004) The clinical features, diagnosis and natural history of nonalcoholic fatty liver disease. Clin Liver Dis 8:521-533, viii
83 Marrero JA, Fontana RJ, Barrat A et al (2005) Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. Hepatology 41:707-716
84 Altekruse SF, McGlynn KA, Reichman ME (2009) Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. J Clin Oncol 27:1485-1491
85 Greten TF, Papendorf F, Bleck JS et al (2005) Survival rate in patients with hepatocellular carcinoma: a retrospective analysis of 389 patients. Br J Cancer 92:1862-1868
86 Lang H, Sotiropoulos GC, Brokalaki EI et al (2007) Survival and recurrence rates after resection for hepatocellular carcinoma in noncirrhotic livers. J Am Coll Surg 205:27-36
87 Perry JF, Charlton B, Koorey DJ et al (2007) Outcome of patients with hepatocellular carcinoma referred to a tertiary centre with availability of multiple treatment options including cadaveric liver transplantation. Liver Int 27:1240-1248
88 Feldstein AE, Wieckowska A, Lopez AR, Liu YC, Zein NN, McCullough AJ (2009) Cytokeratin-18 fragment levels as noninvasive biomarkers for nonalcoholic steatohepatitis: a multicenter validation study. Hepatology 50:1072-1078
89 Angulo P, Hui JM, Marchesini G et al (2007) The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 45:846-854
90 Friedrich-Rust M, Ong MF, Martens S et al (2008) Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology 134:960-974
91 Kwok R, Tse YK, Wong GL et al (2014) Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease--the role of transient elastography and plasma cytokeratin-18 fragments. Aliment Pharmacol Ther 39:254-269
92 Friedrich-Rust M, Nierhoff J, Lupsor M et al (2012) Performance of Acoustic Radiation Force Impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. J Viral Hepat 19:e212-219
93 Nierhoff J, Chavez Ortiz AA, Herrmann E, Zeuzem S, Friedrich-Rust M (2013) The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: a meta-analysis. Eur Radiol 23:3040-3053
94 Castera L, Foucher J, Bernard PH et al (2010) Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology 51:828-835
95 Murtagh J, Foerster V (2006) Transient elastography (FibroScan) for non-invasive assessment of liver fibrosis. Issues Emerg Health Technol:1-4

96 Carlson JJ, Kowdley KV, Sullivan SD, Ramsey SD, Veenstra DL (2009) An evaluation of the potential cost-effectiveness of non-invasive testing strategies in the diagnosis of significant liver fibrosis. J Gastroenterol Hepatol 24:786-791

97 DeKoven, M. Cost comparison: liver biopsy versus abdominal MRI. Memo to GE Healthcare from IMS Health Incorporated 2010.

98 Myers RP, Fong A, Shaheen AA (2008) Utilization rates, complications and costs of percutaneous liver biopsy: a population-based study including 4275 biopsies. Liver Int 28:705-712

99 Régie de l'assurance maladie du Québec (RAMQ). RAMQ liste de medicaments 2013. Quebec Ministry of Health and Social Services; 2013. Available at https://http://www.prod.ramq.gouv.qc.ca/DPI/PO/Commun/PDF/Liste_Med/Liste_Med/liste_med_cor_2013_07_15_fr.pdf. Accessed August, 2013.

100 Thompson Coon J, Rogers G, Hewson P et al (2008) Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. Br J Cancer 98:1166-1175

101 Aberg F, Maklin S, Rasanen P et al (2011) Cost of a quality-adjusted life year in liver transplantation: the influence of the indication and the model for end-stage liver disease score. Liver Transpl 17:1333-1343

TABLES

Table 1. Model Parameters

Parameters	Base Estimate (Range)	References
Prevalence		
Prevalence of steatosis in general population	0.23 (0.16-0.30)	[4; 5; 60]
Prevalence of steatosis in type 2 diabetes population	0.70	[61]
Prevalence of steatosis in obese population	0.75 (0.64-0.90)	[62] [5; 63; 64]
Prevalence of NASH in general population	0.04 (0.02-0.122)	[2; 23; 65; 66]
Prevalence of NASH in type 2 diabetes population	0.25 (0.25-0.30)	[14; 67]
Prevalence of NASH in obese population	0.20 (0.19-0.50)	[5; 62-64]
Prevalence of NASH-cirrhosis in general population	0.0019 (0.0018-0.0020)	[68; 69]
Prevalence of NASH-cirrhosis in type 2 diabetes population	0.02 (0.02-0.03)	Author's assumptions
Annual transition probabilities		
Probability of developing steatosis	0.029 (0.02-0.04)	[23; 70]
Probability of developing NASH	0.0084 (0.00029-0.088)	[23; 71-73]
Probability of NASH liver-related mortality	0.0038 (0.002-0.01)	[6; 74; 75]
Probability of developing fibrosis	0.089 (0.065-0.092)	[3; 23; 52; 76]
Probability of worsening fibrosis	0.11 (0.10-0.13)	[3; 23; 73; 77]
Probability of developing cirrhosis	0.02-0.06	[6; 78]
Probability of NASH-cirrhosis liver-related mortality	0.034 (0.015-0.049)	[75; 79-81]
Probability of developing decompensated cirrhosis	0.06 (0.04-0.16)	[6; 11; 52; 79; 80]
Probability of decompensated cirrhosis-related mortality	0.16 (0.15-0.38)	[8; 82]

Probability of developing HCC	0.029 (0.017-0.08)	[7; 8; 21; 22; 74; 80]
Probability of hepatoma mortality at year 1	0.52 (0.47-0.58)	[83-85]
Probability of hepatoma mortality in subsequent years	0.068 (0.068-0.23)	[85; 86]
Probability of liver transplantation	0.05 (0.05-0.25)	[21; 87]
Sensitivity for NASH detection		
Plasma cytokeratin-18 fragments	0.77 (0.64-0.92)	[14; 53; 88]
Sensitivity for advanced fibrosis (≥ F3)		
NAFLD fibrosis score	0.64 (0.59-0.70)	[14; 89]
Transient elastography (TE)	0.85 (0.58-0.95)	[14; 15; 50; 90; 91]
Acoustic radiation force impulse (ARFI)	0.89 (0.87-0.99)	[49; 50; 92; 93]
Magnetic resonance elastography (MRE)	0.92 (0.85-0.96)	[26]
Technical failure of elastography methods		
Rate of technical failure of TE	0.16	[94]
Rate of technical failure of ARFI	0.021	[50]
Treatment response		
Histological improvement to lifestyle changes	2.40	[18]
Histological improvement to pioglitazone	1.38 (1.01-1.89)	[19; 20]
Histological improvement to vitamin E	1.35 (0.87-2.09)	[19; 20]
Complications of liver biopsy		
Rate of mortality	0.002	[28]
Rate of major bleeding	0.0065	[28]

Abbreviations: NASH, nonalcoholic steatohepatitis; HCC, hepatocellular carcinoma; NAFLD, nonalcoholic fatty liver disease.

Table 2. Health Care Costs (CAN$, Canadian Dollars)

Parameters	Base Estimate (Range)	References
Annual clinical care costs		
No care	77.20	[32]
Routine care and lifestyle changes	325.00	[32]
Routine care and pioglitazone	2106.20	[32]
Routine care and vitamin E	463.70	[32]
Compensated cirrhosis and pioglitazone	2183.40	[32]
Compensated cirrhosis and vitamin E	540.90	[32]
Decompensated cirrhosis	16,679.50 (10,884-22,475)	[32; 37]
Hepatocellular carcinoma (net over 5 years)	15949.80	[38]
Liver transplant (1st year)	163818.77	[39]
Itemized clinical care		
Specialist consultation (initial)	157.00	[32]
Specialist consultation (follow-up)	105.25	[32]
Primary care doctor consultation	77.20	[32]
Dietitian/counselling	62.75	[32]
Laboratory		
Full blood count	11.03	[32]
Liver function tests	20.70	[32]
Lipids	21.31	[32]
Oral glucose tolerance test	15.68	[32]
Hepatitis C antibody	27.24	[32]
Hepatitis B surface antigen	36.30	[32]
Anti nuclear antibody	27.24	[32]
Screening methods		
NAFLD fibrosis score	12.95	[32]
Plasma cytokeratin-18 fragments (CK-18)	6.44	[36]
Transient elastography (TE)	99.44	[33; 35]
Ultrasound-based elastography (ARFI)	114.62	[33; 35]
Diagnostic method		
Magnetic resonance elastography	333.98 (250-400)	[34; 35]
Liver biopsy	595.60 (450-1300)	[95-97]
Complications		
Post-biopsy complication requiring hospitalization	4579	[98]
Treatment (yearly)		
Pioglitazone (Actos)	1084.05	[99]
Vitamin E (800 IU)	138.7	[99]

Abbreviations: NAFLD, nonalcoholic fatty liver disease.

Table 3. Health-Related Quality of Life

Parameters	Base Estimate (Range)	References
Well	1	Authors' assumption
Steatosis	1.0 (0.86-1)	[41], Author's assumption
NASH	0.85 (0.84-0.86)	[41]
Fibrosis	0.84 (0.83-0.85)	[41]
Cirrhosis	0.80 (0.65-0.89)	[21; 43; 44; 46]
Decompensated cirrhosis	0.60 (0.46-0.81)	[21; 43; 44; 46]
Hepatoma	0.73 (0.50-0.80)	[43]
Surgical resection (1st month)	0.73 (0.62-0.84)	[100]
Liver transplant (1st year)	0.69 (0.62-0.86)	[44; 45; 101]
Liver transplant (after transplant)	0.80 (0.79-0.83)	[101]

Abbreviations: NASH, nonalcoholic steatohepatitis.

Table 4. Base Case Analysis of Costs (CAN$) and Utilities of NASH Screening Strategies for General, High Risk Obese, and High Risk Type 2 Diabetes Populations.

Population Type	Screening Strategies	Cost (CAN$)	QALYs	Incremental Cost (CAN$)	Incremental Benefits (QALYs)	$/QALY (ICER)
General Population	No Surveillance	$6,561	42.0422	—	—	—
	Screening* with MRE confirmation and vitamin E treatment	$9,697	42.1622	$3,136	0.1200	$26,143
	Screening* with MRE confirmation and pioglitazone treatment	$10,563	42.1665	$866	0.0043	$199,870
Obese Population	No Surveillance	$13,703	38.7285	—	—	—
	Screening* with MRE confirmation and vitamin E treatment	$17,197	39.1145	$3,494	0.3861	$9,051
	Screening* with MRE confirmation and pioglitazone treatment	$19,809	39.1289	$2,613	0.0143	$182,364
Type 2 Diabetes Population	No Surveillance	$15,049	38.1394	—	—	—
	Screening* with MRE confirmation and vitamin E treatment	$18,608	38.5848	$3,559	0.4454	$7,991
	Screening* with MRE confirmation and pioglitazone treatment	$21,576	38.6015	$2,968	0.0167	$178,210

*Screening algorithm involving NAFLD fibrosis score and transient elastography for fibrosis detection, and cytokeratin-18 for NASH detection.

Abbreviations: NASH, nonalcoholic steatohepatitis; QALY; quality-adjusted life year; MRE, magnetic resonance elastography.

Figure 1

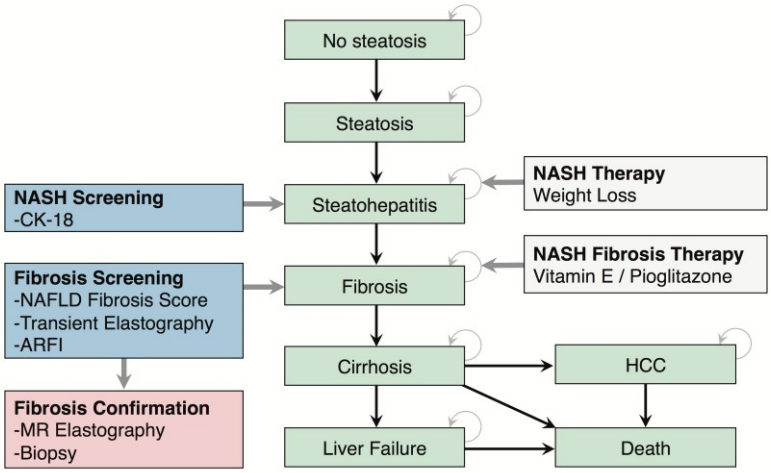

No steatosis
Steatosis
Steatohepatitis
NASH Therapy
Weight Loss
NASH Screening
-CK-18
Fibrosis
NASH Fibrosis Therapy
Vitamin E / Pioglitazone
Fibrosis Screening
-NAFLD Fibrosis Score
-Transient Elastography
-ARFI
Cirrhosis
HCC
Fibrosis Confirmation
-MR Elastography
-Biopsy
Liver Failure
Death

Figure 2A

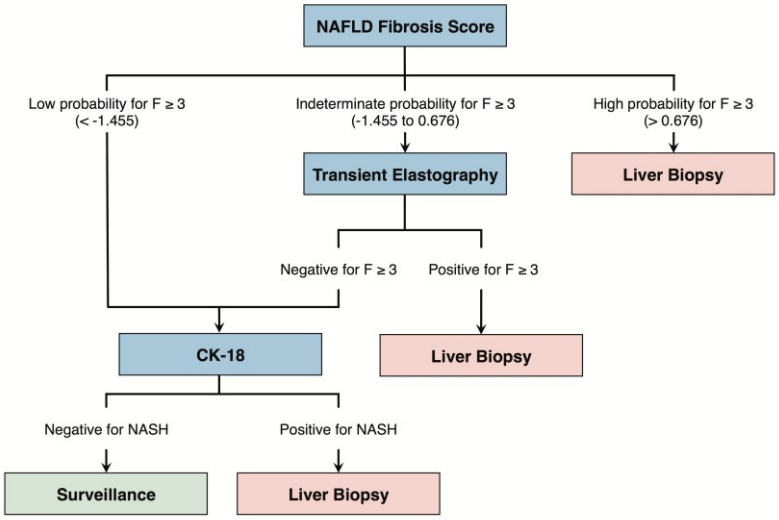

NAFLD Fibrosis Score
Low probability for F ≥ 3
(< -1.455)
Indeterminate probability for F ≥ 3
(-1.455 to 0.676)
High probability for F ≥ 3
(> 0.676)
Transient Elastography
Liver Biopsy
Negative for F ≥ 3
Positive for F ≥ 3
CK-18
Liver Biopsy
Negative for NASH
Positive for NASH
Surveillance
Liver Biopsy

Figure 2B

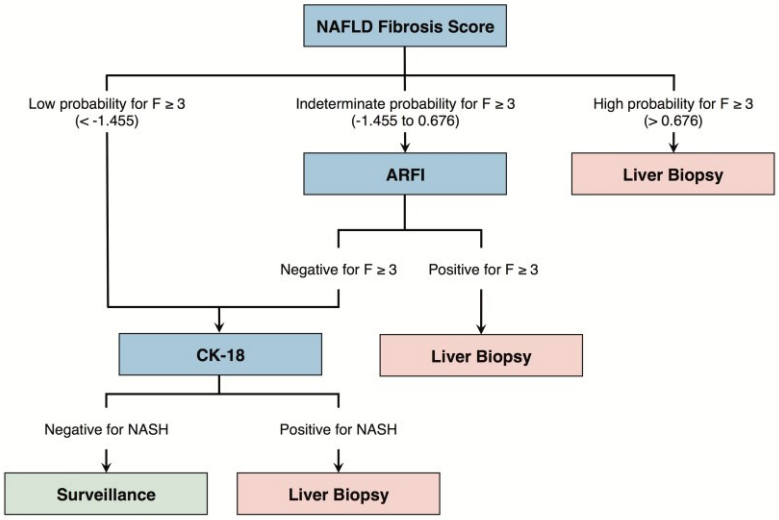

NAFLD Fibrosis Score
Low probability for F ≥ 3
(< -1.455)
Indeterminate probability for F ≥ 3
(-1.455 to 0.676)
High probability for F ≥ 3
(> 0.676)
ARFI
Liver Biopsy
Negative for F ≥ 3
Positive for F ≥ 3
CK-18
Liver Biopsy
Negative for NASH
Positive for NASH
Surveillance
Liver Biopsy

Figure 2C

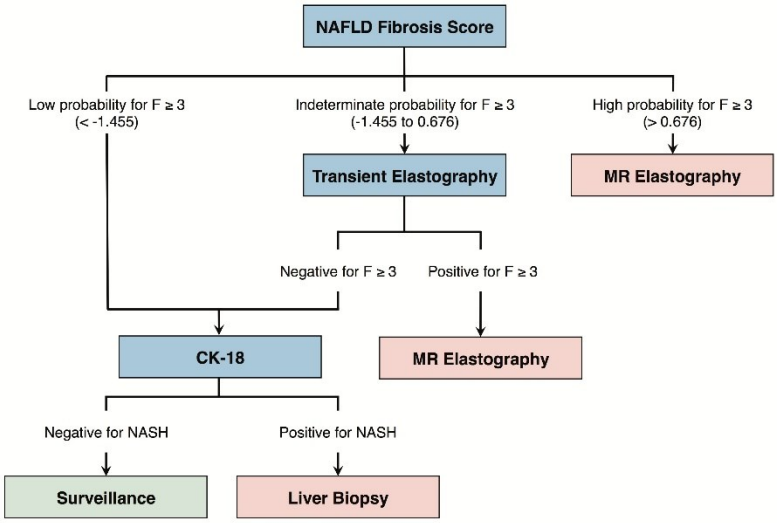
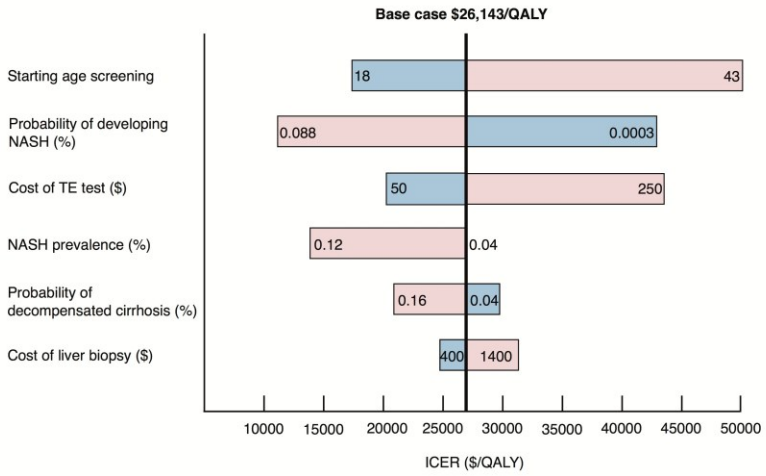

NAFLD Fibrosis Score
Low probability for F ≥ 3
(< -1.455)
Indeterminate probability for F ≥ 3
(-1.455 to 0.676)
High probability for F ≥ 3
(> 0.676)
Transient Elastography
MR Elastography
Negative for F ≥ 3
Positive for F ≥ 3
CK-18
MR Elastography
Negative for NASH
Positive for NASH
Surveillance
Liver Biopsy

Figure 3


Base case $26,143/QALY
Starting age screening
18
43
Probability of developing NASH (%)
0.088
0.0003
Cost of TE test ($)
50
250
NASH prevalence (%)
0.12
0.04
Probability of decompensated cirrhosis (%)
0.16
0.04
Cost of liver biopsy ($)
400
1400
10000
15000
20000
25000
30000
35000
40000
45000
50000
ICER ($/QALY)