Forecasted 2040 Global Prevalence of NAFLD using Hierarchical Bayesian Approach

Authors: Michael H. Le1,2, Yee Hui Yeo1,3, Biyao Zou1,4, Scott Barnet1, Linda Henry1, Ramsey Cheung1,5, Mindie H. Nguyen1,4

Institutions:
1. Division of Gastroenterology and Hepatology, Stanford University Medical Center, Palo Alto, CA
2. Larner College of Medicine, University of Vermont, Burlington, VT
3. Division of Gastroenterology and Hepatology, Cedars Sinai Medical Center, Beverly Hills, CA
4. Department of Epidemiology and Population Health, Stanford University Medical Center, Palo Alto, CA
5. Division of Gastroenterology and Hepatology, Palo Alto Veterans Affairs Medical Center, Palo Alto, CA

Short Title: 2040 NAFLD Forecast

Corresponding Author:
Mindie H. Nguyen, MD, MAS, AGAF, FAASLD
Professor of Medicine and, by courtesy, of Epidemiology and Population Health
Division of Gastroenterology and Hepatology
Stanford University Medical Center
Palo Alto, CA
Email: mindiehn@stanford.edu

Word Count:
Tables: 1
Figures: 4

**Funding:** No external funding to disclose

**Conflicts of interest:**
MHN: Research funding: Pfizer, Enanta, Gilead, CurveBio, Exact Sciences, Helio Health, Glycotest, National Cancer Institute, B.K. Kee Foundation, Vir Biotech; Consulting: Gilead, Intercept, GSK, Exact Science, Novartis, Janssen, Bayer
RC: Research funding: Gilead, Siemens Healthineer

**Author Contributions:**
Study design: MHL, YHY, MHN
Data analysis: MHL, SB, BY, YHY, MHN
Data collection: MHL
Drafting of manuscript: MHL, YHY, LH, MHN
Data interpretation, review and revision of manuscript: all authors

**Abbreviations:**
NAFLD- nonalcoholic fatty liver disease
NASH- nonalcoholic steatohepatitis
HCC- hepatocellular carcinoma
Abstract (word count=247)

Background: Due to increases in obesity and type 2 diabetes, nonalcoholic fatty liver disease (NAFLD) is increasing. Current forecast models may not include non-obese NAFLD. We used a Bayesian approach forecasting the prevalence of NAFLD through 2040.

Methods: Prevalence data from 245 articles involving 2,699,627 persons were used with a hierarchical Bayesian approach to forecast the prevalence of NAFLD through 2040. Subgroup analyses were conducted for age, gender, presence of metabolic syndrome, region, and smoking. Sensitivity analysis was conducted for clinical setting and study quality.

Results: Forecasted 2040 prevalence rate was 55.7%, a 3-times increase since 1990 and 43.2% increase from the 2020 prevalence of 38.9%. The estimated average yearly increase since 2020 was 2.16%. For those aged <50 years and ≥50 years old, the 2040 prevalence rates were not different (56.7% vs 61.5%, P=0.52). There was a significant difference in 2040 prevalence by sex (males- 60% vs 50%, P+) but trend is steeper for females (2.5% vs 1.5%, P=0.025). No difference in trends over time by region (P=0.48). The rate of increase was significantly higher in those without metabolic syndrome (3.8% vs 0.84%, P=0.003) and for smokers (1.4% vs 1.1%, P=0.011). There was no difference by clinical/community setting (P=0.491) or the quality of the studies (P=0.85)

Conclusion: By 2040, over half the adult population is forecasted to have NAFLD. The largest increases occur in women, smokers and those without metabolic syndrome. Intensification of efforts raising awareness and determining long term solutions addressing driving factors of NAFLD are needed.
**Key Words** - women, metabolic syndrome, smoking, trends
**Introduction**

Nonalcoholic fatty liver disease (NAFLD) is a metabolically related liver disease and is strongly associated with the metabolic conditions of insulin resistance, type 2 diabetes and obesity. [1] These three conditions alter the storage of fat in the liver which can lead to hepatic steatosis. In NAFLD, steatosis occurs in greater than 5% of hepatocytes in the absence of excessive alcohol use, viral hepatitis and steatogenic drugs. Currently, the global prevalence of NAFLD in adults is 25%-30% though the prevalence rates vary by country with the highest rates found in Middle East and North Africa (MENA) regions and South America followed by Asia, North America, Europe and then Africa. [2,3,4] NAFLD is also prevalent in children with the prevalence rate estimated to 7.40% in the general pediatric population. [2]

NAFLD can be a progressive disease leading to nonalcoholic steatosis (NASH), fibrosis, cirrhosis, hepatocellular carcinoma, liver transplantation and patients are at an increased risk for mortality especially cardiovascular mortality. [2, 5-7] Although only approximately 20% of those with simple fatty liver (NAFL) will progress in their disease state, the large number of the general population affected can lead to a large increase in health care resource utilization.[6]. In fact, economic studies on NAFLD show that those with NAFLD have increased health care costs and increased years of disability. [8-10]. Those with NAFLD have also reported decreased quality of life especially in their ability to be physically active. [11]

While work continues on how best to diagnose NAFLD and NASH as well as other treatments outside life style modification, gaining an appreciation of the forecasted growth of NAFLD if the trends of obesity and type 2 diabetes mellitus continue to grow is important for policy decision makers and healthcare providers. [12]
A few recent studies have forecasted NAFLD prevalence up to 2030. [13-17] These studies have suggested that since NAFLD appears to be increasing at the same rate as the epidemics of obesity and diabetes such that if the prevalence of obesity and diabetes level off through 2030, there will be only a modest increase in the total number of NAFLD cases (increases of 0 to 30%). Europe and Germany currently have one of the highest prevalence of NAFLD but by 2030, Italy is expected to have a prevalence of 29.5% outpacing Germany while France is expected to have the lowest prevalence rate in Europe. In Asia, China is expected to have the largest increase in prevalence of NAFLD due to its rapid urbanization and was forecasted to reach a prevalence of 28.4% by 2030. [13-17] However, using data on NAFLD prevalence from 167 studies and almost 1.5 million people, a systematic review and meta-analysis of NAFLD prevalence for China published in 2020 already found a NAFLD prevalence of 29.9% which was still increasing. [18]

In addition, most of these studies were either country-specific and/or based on population obesity prevalence, rather than published data for past trends of NAFLD prevalence. Markov modeling was also used as the basis for the analysis which has inherent biases in that people can only be in one stage at a time, many simulations must be run to account for the multiple stages of disease and the inputs are based on transition probabilities that are only in their infancy of development at this time.[19] Therefore, to overcome these limitations, we have forecasted future prevalence using Bayesian modeling based on published data of past NAFLD prevalence trends.

**Methods**

**Data source**

We utilized prevalence data from 245 articles involving 2699627 persons that were identified through a systematic literature search review of PubMed, Embase, and the Cochrane Library from inception to March 2020. The details of the systematic review and characteristics of these 245 articles are previously described in a prior published study that evaluated the current and past trend of NAFLD prevalence [2]
Briefly, NAFLD was diagnosed via liver ultrasound in the vast majority of the studies (202 studies, 82.4%) and via other imaging methods such as computed tomography (7 studies), magnetic resonance imaging (3 studies), controlled attenuation parameter (4 studies), fatty liver index (22 studies), liver biopsy (1 study) or a combination of methods (5 studies). Only data from ultrasound-based studies were used for forecasting analysis except for sub-analysis by geographic region due to limited data for some regions as further described below.[2]

Aggregated subgroup data are available for forecasting analysis for the following subgroups. By age, there are 101 studies providing data for 1,615,429 patients aged <50 and 72 studies providing data for 403,237 patients aged 50 or older. By sex, there are 148 studies providing stratified data for 899,753 male persons and 789,004 female persons. By geographic region, there are 6 studies (23,952 patients) from North America, 11 studies (15,062 patients) from Europe, 182 studies (2,385,999 patients) from Asia, 2 studies from Africa, and 3 studies from South America. Therefore, forecasting analyses were only performed for North America, Europe, and Asia. Additionally, 37 studies provided stratified data for 54,736 patients with metabolic syndrome (metabolic syndrome) and 198,432 patients who did not have metabolic syndrome, and 62 studies provided data for 230,188 smokers and 741,399 non-smokers. Lastly, there were 97 clinical center studies (143,480 patients), 105 community-based studies (1,434,840 patients), 127 studies (1,241,714 patients) with quality scores (based on a modified Newcastle Ottawa scale [5]. ≤7 and 79 studies (1,184,146 patients) with quality scores >7 to provide data for sensitivity analyses by study setting and study quality.

**Statistical analysis**

*Forecast prediction model*

We utilized a hierarchical Bayesian approach to best describe and fit the prevalence of NAFLD. A random effects parameter was used to address heterogeneity across the prevalence studies with NAFLD prevalence as included in a prior study. [2] Individual study prevalence of NAFLD overall was considered
as the proportion of subjects reported as having NAFLD divided by the total number of reported subjects. When using Bayesian models, prevalence may be considered a random variable with an unknown parameter. Therefore, prevalence estimates of NAFLD were transformed using a logit transformation. Logit transformation follows a normal distribution: \( \text{logit}(p[i]) = \mu_i \) and \( \mu_i \sim \text{Normal} (\text{Mu}, \text{precision } \mu) \).

The base model to pool the overall prevalence of age-related NAFLD was: \( u_i = \beta_0 + \beta^*\text{[study year - mean-year]} \), where mean-year was centered by the overall mean year across all studies. Lastly, a random effects parameter due to study heterogeneity was included using the conjugate gamma distribution: Gamma (0.01, 0.01). The Gamma distribution is utilized for unknown quantities for values between 0 and infinity. All coefficients and the intercept in the model were specified with non-informative normal priors—ie, \( \beta \sim \text{Normal} (0, 1 \times 10^{-6}) \), \( \text{Mu} \sim \text{Normal} (0, 1 \times 10^{-6}) \). We used the Just Another Gibbs Sampler (R2JAGS) and Markov-chain Monte Carlo procedure (MCMCpack) within RStudio (R 3.6.3, Boston MA) to estimate posterior distributions. Convergence estimation was assessed by Gelman–Rubin convergence statistics.

**Subgroup analysis**

In addition to overall global NAFLD prevalence forecast to 2040, we performed forecasting for subgroups where data were available. In total, stratified analyses were performed for age (≤50 vs. ≥50 years), sex (male vs. female), geographic region (North America, Europe, and Asia), by the presence of metabolic syndrome (with metabolic syndrome vs. no metabolic syndrome), and by the presence of smoking (smoking history vs. no smoking).

**Sensitivity analysis**

Finally, we performed sensitivity analysis to evaluate for potential differences of forecasted prevalence by the type of study setting (clinical center vs community based) and by study quality using scores based on the Newcastle Ottawa scale (≤7 vs. >7).
Results

Overall global forecasted NAFLD prevalence

Figure 1 displays the overall forecast for NAFLD through 2040. As seen, the prevalence of NAFLD in 1990 was 18.2% with a forecasted prevalence rate in 2040 to be 55.7%, a 3-fold increase from 1990 and 43.2% increase from the 2020 prevalence of 38.9%. The estimated average yearly increase since 2020 is 2.16% (Table 1).

Forecasted NAFLD prevalence for subgroups

By age and sex

We found that the trends for the prevalence of NAFLD increasing for those aged <50 years old and those ≥50 years old were similar (Figure 2A, P=0.52). For those <50 years old, the prevalence of NAFLD is expected to be 56.7% by 2040 while those who are ≥50 years old, the prevalence is expected to be greater than 60%. The average annual percent change of NAFLD prevalence since 2020 are 2.44% and 3.46% for people <50 and ≥50 years of age, respectively (Table 1).

On the other hand, there was a significant difference in the trends for the increase of the prevalence of NAFLD by sex whereby the females have a larger spike in their trend noted to begin around the year 2025 (Figure 2B, P=0.025), though the prevalence increases over time for both males and females. Males are forecasted to have a NAFLD prevalence of almost 60% by 2040 and a yearly increase of 1.48% from 2020 when it was 45.5%. Females NAFLD prevalence in 2040 is forecasted to be approximately 50% and a yearly increase of 2.52% from 2020 when it was approximately 33.3% (Table 1).

By geographic region

When we investigated the forecast of NAFLD by region, we found there was no difference in increasing trends overtime by region (Figure 3, P=0.48). The prevalence of NAFLD in Europe is forecasted to increase from approximately 18% in 1990, 43.4% in 2020 to over 60% by 2040 with Asia noting an
almost identical increase both of which is an average yearly increase of 2.19% and 2.70%, respectively (Table 1). In North America the prevalence increase from approximately 30% in 1990 to 43.1% in 2020 and to 50% by 2040 with an average yearly increase of 1.01% since 2020.

By the presence of metabolic syndrome or cigarette smoking

There was also a significant difference in the increasing trends of NAFLD by those with and without metabolic syndrome (Figure 4A, \(P=0.003\)). However, the rate of increase is significantly higher for those without metabolic syndrome compared to those with metabolic syndrome with the average annual increase of 3.78% and 0.84%, respectively (Table 1).

There was also a significant difference in the increasing trend of NAFLD prevalence by smoking status (Figure 4B, \(P=0.011\)). Those who were smokers saw an increase of NAFLD prevalence from 28.1% in 1990 and 49.2% in 2020 to 63% by 2040 whereas those who were non-smokers saw an increase from 25.7% in 1990 and 35.5 in 2020 to 43.5% in 2040. The average rate of change for smokers was 1.40% while for non-smokers it was only 1.13% (Table 1).

Sensitivity analyses

There was no difference in the increasing trend of NAFLD prevalence by clinical center or community setting (\(P=0.491\)) or the quality of the studies used to generate forecast data (\(P=0.85\)) (Supplementary Figures 1A and B).

Discussion

In this study we used past data and trend to forecast the prevalence of NAFLD by 2040. This method of forecasting is different than what has recently been completed and reported where investigators used Markov modeling, and we also performed the forecast based on published NAFLD prevalence data rather than surrogate data on obesity.[13-17] Though we recognize the increasing burdens of obesity and
diabetes are driving the increase in NAFLD, not all patients with NAFLD have diabetes or obesity, [20-22] so our modeling methods takes this information into account by using NAFLD historic data that includes trend data from patients with NAFLD regardless of the presence of obesity, rather than relying mainly on trends of obesity. We found that if the current increase in NAFLD continues at the same pace, by 2040, over half the adult population will have NAFLD.

The projection from the current study is much higher than previous estimates which forecasted the prevalence of NAFLD to be only up 33% by 2030.[13-17] In fact, we found in this analysis that the estimated NAFLD prevalence of 38.9% for 2020 is almost identical to two recent prevalence estimates of 37.3% and 37.6% from two recent meta-analyses, [2,23] and all of these estimates are higher than the highest forecasted estimate of 33% forecasted by an earlier study[13-17] As such, our methodology appears robust and that considering the risk of adverse outcomes to include cirrhosis, liver transplantation, and mortality, the burden of NAFLD can become enormous if it continues to increase at our projected rate of 2.16% a year unless aggressive action is taken to reverse this trend. In fact, NASH is now poised to become the leading indicator for liver transplantation in the next few years as well as NASH HCC is already one of the leading indicators for liver transplantation among those with HCC. [24,25]

There are several unique findings of this study which may help policy makers in making decisions as to where interventions may be most warranted and to increase awareness of NAFLD. [26-29] When we forecasted trends by age, there was no difference in prevalence rates in 2040 by those < 50 or ≥ 50. Both groups will have prevalence rates at almost 50% and these results suggest that NAFLD should no longer be considered just a disease found in older individuals. In fact, studies are showing that rates of NAFLD are increasing in children as well. [5,30,31] The global NAFLD prevalence among children is forecasted to be 30.7% by 2040. [5] Thus, treatment interventions will need to be appropriate for all age groups and early intervention at early age would be crucial to bring down the prevalence in the adult population.
Our analysis by sex also provided interesting results. Both males and females will incur an increase in their forecasted NAFLD prevalence in 2040 but the female will increase at a higher rate. Although NAFLD is currently more prevalent in males, NASH has been found to be more prevalent in females. [32,33] In fact, NASH is the leading indicator for liver transplant among women. [34] These findings are worrisome as they could suggest that females may be at higher risk for adverse outcomes in the coming years. Future research is needed to determine the significance of this finding.

We found those without metabolic syndrome will have a significantly higher rate of increase compared to those with metabolic syndrome. This finding may seem counterintuitive since NAFLD is a metabolic based liver disease, but as our understanding of NALFD increases, investigators have recently suggested that the main driver for NAFLD development is the presence of insulin resistance and the related liver lipotoxicity.[34,35] In addition, it appears that visceral obesity is a better indicator of the status of obesity rather than body mass index (BMI) and waist circumference or abdominal adiposity correlated with risk of mortality rather than metabolic syndrome itself.[36]

There was no difference in NAFLD prevalence over time by country or region. All studied areas are expected to have increased prevalence. However, we did note that those who are smokers would have a significantly higher rate of NAFLD in 2040 than those who are not smokers. As such, Asia and Eastern Europe may need to watch this trend more closely as they are known to have one of the highest rates for smoking whereas in the United States and other western nations, smoking is decreasing. [37,38]

The strength of the study is the use of Bayesian statistical modeling, large sample size and widespread geographic coverage to forecast a more accurate assessment of NAFLD trends. Forecasting, however, does not address the impact of those trends nor do forecasting models take into consideration any changes in factors accounting for observed trends. For example, if the observed trajectory of the causative factors of NAFLD were to change, we would be unable to forecast this change. For example, the advances in
weight-reduction and antidiabetic medications may reduce the obesity and metabolic disease burden that may reduce the prevalence and incidence of NAFLD, but these new and often expensive medications are likely not accessible to people in most regions of the world and many disadvantaged populations residing in high-income countries. At the same time, the current COVID-19 pandemic has caused severe disruption in access to healthcare and rising living cost that may worsen the current obesity and metabolic disease epidemics which can contribute to even higher NAFLD prevalence. In fact, population-based data from the U.S. have shown increased mortality during the pandemic as compared to predicted level based on past trends for several chronic liver disease including NAFLD and especially in ethnic minorities. [39-40] However, the information provided in our analysis provides a worst-case scenario which may help policy makers in developing policies to raise awareness of NAFLD while simultaneously providing direction and funding for reversing the upward trajectory for both obesity and type 2 diabetes mellitus. This is particularly relevant as increasingly NAFLD is recognized as a metabolic liver disease with a proposed name change to metabolic fatty liver disease (MAFLD), [41] and therapeutic strategies should include not only liver specific agents such as anti-fibrotics but also systemic metabolic modifying agents such as antidiabetic and antilipid medications. [42,43]

**Conclusion**

In this forecasting study, we project that the global NAFLD prevalence in 2040 will be 55.2% with the highest prevalence of over 60% forecasted for Asia and Europe. In addition, women, smokers and those without metabolic syndrome are forecasted to have the largest increases in prevalence. Given that approximately 20% of those with NAFLD can progress to more advanced disease and that the presence of NAFLD increases the risk of mortality, efforts may need to intensify to not only raise awareness of NAFLD but also to determine long term solutions to the driving factors of NAFLD such as obesity, insulin resistance and type 2 diabetes mellitus.
References


Table 1: Annual rate of change in the prevalence of NAFLD between 2020 to 2040, overall and in subgroups

<table>
<thead>
<tr>
<th>Patient groups</th>
<th>Annual rate of change in NAFLD prevalence (%)</th>
<th>95% Confidence interval (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2.16</td>
<td>1.60-2.71</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>2.44</td>
<td>1.57-3.31</td>
</tr>
<tr>
<td>&gt;50</td>
<td>3.46</td>
<td>1.05-3.46</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.48</td>
<td>1.08-1.89</td>
</tr>
<tr>
<td>Female</td>
<td>2.52</td>
<td>1.56-3.49</td>
</tr>
<tr>
<td>Geographic location*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America</td>
<td>1.01</td>
<td>0.00-3.02</td>
</tr>
<tr>
<td>Europe</td>
<td>2.19</td>
<td>0.12-4.25</td>
</tr>
<tr>
<td>Asia</td>
<td>2.70</td>
<td>2.23-3.17</td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>3.78</td>
<td>2.90-4.66</td>
</tr>
<tr>
<td>Yes</td>
<td>0.84</td>
<td>0.74-0.94</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.13</td>
<td>0.00-2.71</td>
</tr>
<tr>
<td>Yes</td>
<td>1.40</td>
<td>0.83-1.97</td>
</tr>
</tbody>
</table>

NAFLD, non-alcoholic fatty liver disease

*Data insufficient for subgroup analysis for Africa and South America
Figure 1. Overall global forecasted NAFLD prevalence
Figure 2: Forecasted NAFLD prevalence by age and sex. *P values compares rate of increase between subgroups via linear regression.

n=number of studies
Figure 3: Forecasted NAFLD prevalence# for North America, Europe, and Asia

#Based on past trend of ultrasound diagnosed prevalence of NAFLD. North America subgroup includes studies that used FLI to diagnose NAFLD. Data were insufficient to perform forecasting for Africa and South America.

*P values compares rate of increase between subgroups via linear regression

n=number of studies
Figure 4: Forecasted NAFLD prevalence by the presence of metabolic syndrome or cigarette smoking

*P values compares rate of increase between subgroups via linear regression

n=number of studies
Supplemental Figure 1: Sensitivity analyses of forecasted NAFLD prevalence by type of study setting and by study quality

*P values compares rate of increase between subgroups via linear regression

n=number of studies