Modeling the Epidemic of Nonalcoholic Fatty Liver Disease Demonstrates an Exponential Increase in Burden of Disease – Appendix

Table of Contents

Section 1: Model Description ....................................................................................................... 2
Section 2: Progression Rates ...................................................................................................... 3
Section 3: Back Calculation of New NAFLD Cases (Incidence) .................................................. 6
Section 4: Relative Incidence ....................................................................................................... 8
Section 5: Delphi Process ........................................................................................................... 11
Section 6: Supplementary Reference List .................................................................................. 12
Section 1: Model Description

The nonalcoholic fatty liver disease (NAFLD) Markov model was designed using Microsoft Excel® 2007 (Microsoft Corp., Redmond, WA) to track the NAFLD population by fibrosis stage as well as nonalcoholic steatohepatitis (NASH) status from 1950-2050. The relative impact of incident NAFLD cases occurring prior to 1950 was negligible and was not included in the analysis. Model generated uncertainty intervals (UI) were calculated using high/low Beta-PERT distributions around inputs and conducting Monte Carlo analysis using Oracle Crystal Ball® (Oracle Corp., Redwood City, CA, Release 11.1.3708.0).

The model begins with the annual estimated number of incident NAFLD cases (as shown in the figure below). Fibrosis progression of all cases was modeled over time. The number of cases at any stage of liver disease was calculated and tracked by age and gender. The population was aged in one year age cohorts through age 84 and cases aged ≥85 years were tracked as a single cohort. Each year, the population in each age group, except for the ≥85 year cohort, was moved to the next age to simulate aging (after taking into account age specific mortality and new cases). Background population data for the US were obtained from the United Nations’ population database by age, gender and one year age cohort (1).

Disease progression was estimated through fibrosis and liver disease stages (Figure 1) with annual adjustment for mortality (background mortality, and excess cardiovascular deaths), as described below. Liver related deaths were handled separately using published progression rates from advanced liver disease stages (decompensated cirrhosis, hepatocellular carcinoma (HCC)) and liver transplantation. Total cases by disease stage were calculated by summing the total cases from the previous year, the previous age, and new cases; and subtracting mortality and progression to subsequent stages as shown below.

\[
Total_{\text{Cases}}_{\text{Stage} \& \text{Year} \& \text{Age Cohort}_z} = (Total_{\text{Cases}}_{\text{Stage} \& \text{Year}-1 \& \text{Age Cohort}_z-1}) + \text{New Cases}_{\text{Stage} \& \text{Year} \& \text{Age Cohort}_z} - \text{Mortality}_{\text{Stage} \& \text{Year} \& \text{Age Cohort}_z} - \text{Progressed}_{\text{Stage} \& \text{Year} \& \text{Age Cohort}_z}
\]

where:

\[
\text{New Cases}_{\text{Stage} \& \text{Year} \& \text{Age Cohort}_z} = (Total_{\text{Cases}}_{\text{Stage}-1 \& \text{Year}-1 \& \text{Age Cohort}_z})(\text{Progression Rate}_{\text{Stage}-1 \rightarrow \text{Stage} \& \text{Age Cohort}_z})
\]
\[ Mortality_{Stage_x \& Year_y \& Age\ Cohort_z} = \] 
\[ (Total\ Cases_{Stage_x \& Year_{y-1} \& Age\ Cohort_z})(Adjusted\ Mortality\ Rate_{Age\ Cohort_z}) \]

\[ Progressed_{Stage_x \& Year_y \& Age\ Cohort_z} = \] 
\[ (Total\ Cases_{Stage_{x-1} \& Year_{y-1} \& Age\ Cohort_z})(Progression\ Rate_{Stage_x \rightarrow Stage_{x+1} \& Age\ Cohort_z}) \]

**Figure 1. NAFLD Markov Model**

**Section 2: Progression Rates**

Disease progression was simulated by multiplying the total number of cases at a particular stage of the disease by a progression rate to the next stage (Figure 1) and validated using national surveillance data for incidence of NAFLD-related HCC, as described below. Age specific fibrosis progression rates were back-calculated based on assumptions for the distribution of cases by NASH status and fibrosis stage, as described below. In addition, the reported numbers of NAFLD related HCC cases, HCC deaths and decompensated cirrhosis cases were available for the US, and fibrosis progression rates were modified to fit reported data.
Table 1). For the purpose of the model, progression rates were assumed to be the sum of forward progression minus the rate of regression, which is common among NAFLD cases based on studies of consecutive liver biopsies (2).

Table 1. Disease Stage Transitions

<table>
<thead>
<tr>
<th>Disease Stage Transition</th>
<th>Model Transition Rate</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0 to F1</td>
<td>0.68% – 2.16%</td>
<td>Back-calculated</td>
</tr>
<tr>
<td>F0 to HCC</td>
<td>0.00035%</td>
<td>(3, 4)</td>
</tr>
<tr>
<td>F1 to F2</td>
<td>4.18% – 13.25%</td>
<td>Back-calculated</td>
</tr>
<tr>
<td>F1 to HCC</td>
<td>0.0093%</td>
<td>(3, 4)</td>
</tr>
<tr>
<td>F2 to F3</td>
<td>4.18% – 13.25%</td>
<td>Back-calculated</td>
</tr>
<tr>
<td>F2 to HCC</td>
<td>0.019%</td>
<td>(3, 4)</td>
</tr>
<tr>
<td>F3 to F4 (Cirrhosis)</td>
<td>5.06% – 9.90%</td>
<td>Back-calculated</td>
</tr>
<tr>
<td>F3 to HCC</td>
<td>0.038%</td>
<td>(3, 4)</td>
</tr>
<tr>
<td>Cirrhosis to Decomp Cirrhosis</td>
<td>3.79%</td>
<td>(5)</td>
</tr>
<tr>
<td>Cirrhosis to HCC</td>
<td>0.33%</td>
<td>(3, 4)</td>
</tr>
<tr>
<td>Decomp Cirrhosis to Liver Rel. Death</td>
<td>20.0%</td>
<td>(5)</td>
</tr>
<tr>
<td>HCC to Liver Rel. Death (Yr 1)</td>
<td>61.0%</td>
<td>(4)</td>
</tr>
<tr>
<td>HCC to L.R. Death (Sub Yrs)</td>
<td>16.2%</td>
<td>(6)</td>
</tr>
</tbody>
</table>

Rates of development of HCC were calibrated to national data from the Surveillance, Epidemiology, and End Results Program (SEER) for liver and intrahepatic bile duct cancer incidence (7). It was conservatively assumed that 72% of all liver cancers were HCC (8). In addition, it was assumed that 12.6-14.8% of modeled incident HCC cases during 2004-2009 were attributable to NAFLD/NASH based on a study of SEER and Medicare-linked data for 4,929 HCC cases during the same time period (4). Among the NAFLD/NASH-related HCC cases, it was assumed that 64% would occur among cirrhotics (3). Given the model calculated cirrhotic population, the annual transition rate was estimated at 0.33%. The remaining 36% of incident HCC cases occurred among F0-F3 cases. The incidence rate among F3 cases was
back-calculated and progression decreased exponentially with each decreasing level of fibrosis from 0.038% (F3 to HCC) to 0.00035% (F0 to HCC) (Table 1). Using data from the SEER-Medicare linkage study (4), it was found that NAFLD-related HCC cases currently experience higher rates of mortality as compared to HCC attributable to HCV. A first year mortality rate of 61% was applied to incident HCC cases, and subsequent year mortality was estimated based on long-term survival data (6). A long term follow up study of individuals with NASH-related cirrhosis reported that 45% experienced liver failure or decompensated cirrhosis defined as an increase in CTP score of 2 points in patients with Child class A cirrhosis over a twelve year time frame (13). An annual progression rate of 3.8% decompensation among cirrhotics was derived from this study and applied in the model.

Transitions between disease states and liver-related mortality rates were compared between the current model and the analysis by Younossi et al., 2016 (9) (}
Figure 2), which describes a steady state Markov chain model that simulates NAFLD incidence, disease progression and regression, and was calibrated to published literature and validated using DisMod II (10). The number and characterization of disease states differed between the models, however both models were calibrated to published data, including HCC incidence reported by SEER. The Younossi et al., 2016 model included transition rates that were applied directly to the NAFL and NASH populations, while the current model applied fibrosis transition rates, with the distribution of NAFL and NASH calculated in a separate step based on fibrosis progression among the entire NAFLD population.
Figure 2).
Figure 2. Comparison of Model and Younossi et al., 2016 Disease State and Mortality Transition Rates

FB (fibrosis); CC (compensated cirrhosis); DCC (decompensated cirrhosis); LT (liver transplant)

Section 3: Back Calculation of New NAFLD Cases (Incidence)

The annual number of new NAFLD cases (incidence) was not available and was back-calculated using the total number of NAFLD cases. Total NAFLD cases in 2015 (83.1 million) was assumed to be the sum of new NAFLD cases from 1950-2015 after adjustment for mortality as shown in the equation below.

\[
\text{Total NAFLD Cases (Prevalent Cases)}_{\text{Year}_y} = \sum_{t=1950}^{y} \left( \text{New NAFLD Cases (Incident Cases)}_t - \text{Mortality}_t \right)
\]

It was assumed anyone who developed incident NAFLD prior to 1950 is no longer alive. The total number of NAFLD cases in 2015 was known and the annual number of deaths (mortality) was calculated in the model using liver related and non-liver related deaths. Solving the above equation for new NAFLD cases provides the average number of new NAFLD cases per year.
To account for the fact that the number of new NAFLD cases was not constant over time, a relative incidence curve was used (Section 4).

Annual relative incidence values were used to describe changes in the annual number of new NAFLD cases over time. In step 1 of the back calculation, the relative incidence was estimated by year using the change in obesity and diabetes as described in Section 4. The Excel® Solver add-in was used to solve for the constant times the annual relative incidence that resulted in the known prevalence (83.1 million total NAFLD cases in 2015) after adjusting for mortality. This constant multiplied by the relative incidence provided the number of new NAFLD cases per year.

In step 2 of the back calculation, annual incident cases were distributed by age and gender to fit the adjusted NAFLD prevalence (Figure 3). A weighting factor was applied to reported prevalence by age and gender from national surveillance data (11) in order to total to 83.1 million cases in 2015. The percentage of the incident population allocated to each age and gender cohort in years 1950-1965 was set equal to 1966 and trended linearly in 5 five-year increments until 2011, at which point the percent of incident cases allocated to each age and gender cohort were held constant until 2030.
Section 4: Relative Incidence

Relative changes in the number of total NAFLD cases were imputed from data related to trends for adult prevalence of obesity and diabetes mellitus (DM). Distribution of NAFL vs. NASH in these populations was used to impute the trends for these histological phenotypes (11-13).

Data were reported by the Centers for Disease Control and Prevention (CDC) for adult obesity beginning in 1960 when an estimated 13.3% of adults aged 20-74 years were classified as obese (body mass index ≥30) (14). By 2009-2010, the prevalence of adult obesity had increased to 35.7% based on National Health and Nutrition Examination Survey (NHANES) data. As shown in Figure 4, later estimates based on NHANES data estimated adult obesity at 35.1% in the US during 2011-2012, suggesting a leveling off of adult obesity rates during 2003-2012, after decades of steady growth (15). Future obesity prevalence among adults has been projected based on historical trends. Based on one projection scenario, adult obesity would increase from 30.94% in 2010 to 42.19% in 2030 (16).
Unlike obesity, DM data were less readily available in years prior to 1990 due to changes in screening and diagnosis levels. In 1958, CDC estimated that 0.93% of adults had been diagnosed with DM, increasing to a maximum of 7.18% in 2013 (17). Estimates for combined diagnosed and undiagnosed DM were considered beginning in 1988-1994 when CDC estimated adult DM prevalence of 8.4% (18). As shown in Figure 4, by 2007-2010, adult DM prevalence was estimated at 11.4%. Finally, a modeling study considered the future prevalence of DM in the US under different intervention strategies. Without interventions, the prevalence of adult DM was projected to increase 12.9% in 2007 to 22.7% in 2030 (19).

The curves shown in Figure 4 were converted to total adult obese and DM cases by multiplying the prevalence times the US adult population in a given year. Incidence of obesity and DM were calculated by subtracting prevalence in each year from the prior year. The resulting incident cases were scaled as shown in Figure 5 Relative incidence for obesity and DM were used to develop the relative incidence curve for NAFLD, where it was assumed that peak NAFLD incidence would occur after peak adult obesity incidence and before peak DM incidence.
Figure 5. Relative incidence of obesity, DM and NAFLD
## Section 5. Delphi Process

<table>
<thead>
<tr>
<th>Activities</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase 1 – Data Gathering</strong></td>
<td>1a</td>
</tr>
<tr>
<td>Identify country experts who are willing to collaborate</td>
<td></td>
</tr>
<tr>
<td>- Experts were identified through NAFLD-related scientific contributions, or through referrals and recommendations from leading researchers.</td>
<td></td>
</tr>
<tr>
<td>Literature Search</td>
<td>1b</td>
</tr>
<tr>
<td>- Review the internal database for previously identified sources</td>
<td></td>
</tr>
<tr>
<td>- Review online sources (e.g., CDC, etc.) to capture non-indexed sources</td>
<td></td>
</tr>
<tr>
<td>- Run a literature search to identify recent publications</td>
<td></td>
</tr>
<tr>
<td>- Summarize input data available through the literature</td>
<td></td>
</tr>
<tr>
<td>- Gather empirical data for new HCC cases, liver transplants, percent of HCC and transplants due to NAFLD, percent of cases with obesity or DM</td>
<td></td>
</tr>
<tr>
<td>- Build draft model based on published data</td>
<td></td>
</tr>
<tr>
<td>- Schedule meeting with experts</td>
<td></td>
</tr>
<tr>
<td><strong>Phase 2 – Country Meetings and Modeling</strong></td>
<td>2a</td>
</tr>
<tr>
<td>Expert Meeting 1 (2-3 hours)</td>
<td></td>
</tr>
<tr>
<td>- Provide a background on the project, model and methodology</td>
<td></td>
</tr>
<tr>
<td>- Review data identified in Phase 1b and highlight gaps in data</td>
<td></td>
</tr>
<tr>
<td>- Request data in local non-indexed journals, unpublished data and any other available data (e.g., hospital-level data) that can be used to fill the gaps</td>
<td></td>
</tr>
<tr>
<td>- Gain agreement on data sources that can used as for extrapolation when no local data are available</td>
<td></td>
</tr>
<tr>
<td>Follow up with Experts Post Meeting 1</td>
<td>2b</td>
</tr>
<tr>
<td>- Send minutes of the meeting and list of remaining action items to experts</td>
<td></td>
</tr>
<tr>
<td>- Follow up with experts to collect missing data and get copies of publications, government reports and unpublished data (e.g., raw hospital or registry-level data)</td>
<td></td>
</tr>
<tr>
<td>- Analyze raw data and send to experts for approval</td>
<td></td>
</tr>
<tr>
<td>Disease Burden Modeling</td>
<td>2c</td>
</tr>
<tr>
<td>- Populate disease burden model with inputs and calibrate model to empirical data</td>
<td></td>
</tr>
<tr>
<td>- Schedule second meeting</td>
<td></td>
</tr>
<tr>
<td>- Develop a slide deck summarizing all inputs and associated data sources</td>
<td></td>
</tr>
<tr>
<td>- Perform a final check of the model and slide deck and approve internally</td>
<td></td>
</tr>
<tr>
<td>Expert Meeting 2 (2-3 hours)</td>
<td>2d</td>
</tr>
<tr>
<td>- Review all inputs as well as data provided by experts since meeting 1 and results of analyses of any raw data provided</td>
<td></td>
</tr>
<tr>
<td>- Gain agreement on all inputs to be used in the model</td>
<td></td>
</tr>
<tr>
<td>- Update the model using any updated inputs</td>
<td></td>
</tr>
<tr>
<td><strong>Phase 3 – Follow-up Analyses</strong></td>
<td>3a</td>
</tr>
<tr>
<td>Follow-up Analyses</td>
<td></td>
</tr>
<tr>
<td>- Update model as necessary and send results to experts</td>
<td></td>
</tr>
<tr>
<td>- Provide support to address follow-up questions</td>
<td></td>
</tr>
<tr>
<td>- Finalize approved inputs and outputs</td>
<td></td>
</tr>
<tr>
<td>- Update analysis as new information becomes available (e.g., new national studies, updated treatment data)</td>
<td></td>
</tr>
</tbody>
</table>
Section 6: Supplementary Reference List


Yilmaz, Y. (2012). Review article: is non-alcoholic fatty liver disease a spectrum, or are steatosis and non-alcoholic steatohepatitis distinct conditions? Aliment. Pharmacol. Ther., 36(9), 815-823.


References


