Identifying Progressing Patients in an Artificial Intelligence (AI) Based Cohort of Nonalcoholic Steatohepatitis (NASH)

Bandaria J, Boussios C, Donadio G, Starzyk K , Gliklich R | OM1, Inc, Boston, MA, USA

Background

- Nonalcoholic steatohepatitis (NASH) is a frequently progressive subset of nonalcoholic fatty liver disease that can be complicated by cardiovascular disease, cirrhosis, and hepatocellular carcinoma
- Although millions of people are estimated to have NASH in the US, a small fraction of those are readily identifiable in existing health-related data.
- Identification of both probable and confirmed NASH is limited by a frequent lack of symptoms, under-diagnosis, absence of targeted treatments and poor billing code specificity (a specific ICD code was not available until late 2016).
- Despite recent progress in the development of non-invasive and radiological approaches to identify and monitor patients, many patients remain that are not diagnosed and/or documented as having NASH.
- Understanding patients at risk of progression is critical to identifying appropriate candidates for clinical trials and triaging patients for more frequent monitoring.

Objective

- To identify and characterize (a) broad cohort of probable NASH patients and (b) characterize those with evidence of progression to more severe disease.

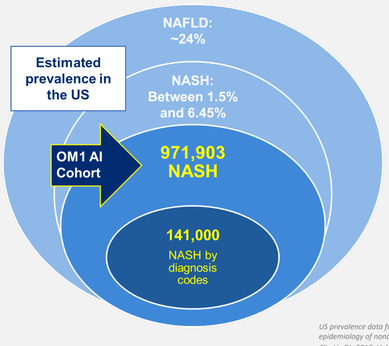
Methods

- The OM1 Real World Data Cloud (OM1, Boston, MA) collects, links and leverages structured and unstructured data from electronic medical records, medical and pharmacy claims and other sources in an ongoing and continuously updating manner. Patient level data are deterministically linked to provide the most comprehensive longitudinal patient journeys for algorithm development
- The data include extensive clinical and claims data on patients seen in a variety of provider practice types across the US, which is important in understanding the multi-systemic burden of the disease and diagnostic pathway
- A combination of sophisticated artificial intelligence (AI) algorithms were used to characterize the likelihood that a patient who is not an ICD10 identified NASH patient is a NASH patient.
- Algorithms were initially applied to a sample of >44 million obese and/or diabetic patients [ROC out-of-sample 0.86], then applied to a larger unselected U.S. population of ~240 million.
- Methods were validated by systematic comparisons to clinically documented NASH patients and alignment of clinical characteristics.
- The OM1 NASH Cohort identified in data available as of May 2019 was explored to characterize patients with by demography and the presence of sever disease indicators.
- FIB-4 scores were calculated using Age (years)×AST (U/L)/[PLT(109/L)×ALT1/2 (U/L)]. Using a lower cutoff value of 1.45, a FIB-4 score <1.45 has a negative predictive value of 90% for advanced fibrosis (i.e., early bridging fibrosis to cirrhosis). FIB-4 >3.25 has a 97% specificity and a positive predictive value of 65% for advanced fibrosis. (Sterling et al. Hepatology 2006;43:1317-1325).

Results

- A total of 971,903 high-likelihood NASH patients were identified, including approximately 15,000 patients under age 18, compared to 141,000 identified by diagnosis code alone (**Figure 1**).
- Geographic distribution of prevalent NASH patients identified by AI is shown in **Figure 2**
- Mean age was 56 years (SD 14.1) and male to female ratio was 51:49 in the AI-identified cohort. Gender and age differences were observed by case identification method and the presence of advanced disease.
- Cirrhosis was documented in 24.4% of patients (including > more than 500 pediatric patients), portal hypertension in 7.8% and history of liver transplant in 1.8% (**Table 1**).
- Of over 237,000 calculated FIB-4 scores in 77,862 patients, 24% of the results were indicative of advanced fibrosis, the majority of which were in patients not identified by related diagnosis codes.

Figure 1. OM1 NASH Cohort



US prevalence data from: Younossi ZM. The epidemiology of nonalcoholic steatohepatitis. Clin Liv Dis 2018; Vol11 (4): 92-94.

Figure 2.

NASH Cohort: Geographic Distribution by State (N= 971,903)

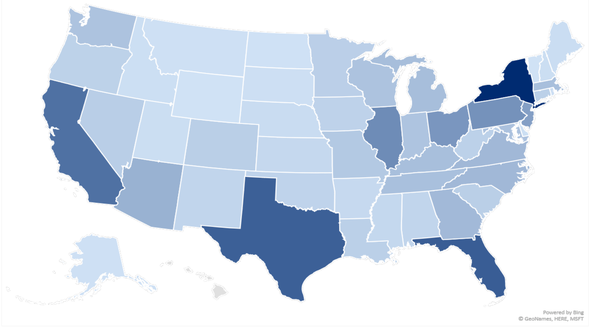


Table 1. Age Distribution by NASH Status

Age group (years)	NASH (AI)	With cirrhosis	With portal hypertension	With liver transplant
Overall	971,903	234,948	75,414	17,577
<18	1.6%	0.2%	0.8%	3.3%
18 to <40	13.1%	4.9%	6.2%	9.2%
40 to <60	44.6%	47.1%	49.0%	45.7%
>=60	40.7%	47.8%	44.0%	41.9%

Diagnoses identified by ICD9 and ICD10 codes, with at least 2 codes at least 30 days apart in any data source (i.e., electronic medical records or medical claims; liver transplant by relevant codes, including history of.

Conclusions

- **Identifying probable cases of NASH, and understanding their early clinical characteristics as well as the potential for progression, has important implications for clinical detection and optimizing patient management, as well as facilitating therapeutic clinical development.**
- **This AI-based approach is a robust and broadly applicable means of identifying probable NASH patients that employs data collected during routine clinical care by the wide variety of healthcare providers managing these patients.**
- **Reliance on biopsy and diagnosis codes alone may not be sufficient to identify patients with progressing NASH in routine practice.**
- **Key applications include exploring potentially aggressive phenotypes (e.g., pediatric NASH, rapidly progressing adult NASH) with the greatest unmet clinical need.**

Presented at 2019 ISPOR EU Annual Conference | 2-6 November 2019 | Copenhagen, Denmark