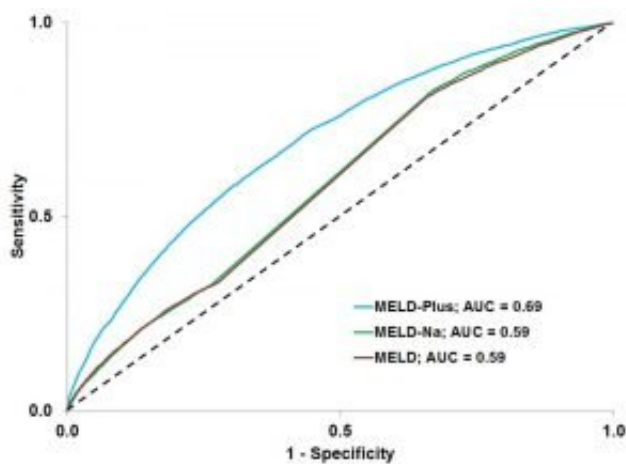


# Machine learning used to surpass MELD in assessing liver disease

October 26 2017, by Uri Kartoun

---



The IBM Explorys Network approximate one-year mortality (24,042 cirrhosis-related admissions). MELD-Plus outperforms traditional MELD and MELD-Na scores.

Before joining IBM Research in Cambridge in September 2016, I completed a three-year postdoctoral training program at Massachusetts General Hospital (MGH), working closely with Harvard faculty. Now, a collaboration between MGH and IBM Research has yielded a satisfying result – our manuscript, "The MELD-Plus: A Generalizable Prediction Risk Score in Cirrhosis," was published in *PLOS ONE*, a peer-reviewed online science journal.

The origins of this collaboration can be traced back nearly two years ago,

when I presented a study at the AMIA Annual Symposium in San Francisco. Right after my talk, the audience had the opportunity to ask questions. IBM Research staff member Kenney Ng asked a question related to a machine learning technique called k-fold cross validation. I was surprised. Typically at such events I am asked to clarify how I conducted our research. I had never been asked such a specific methodological question. I remarked, "You are thinking like a computer scientist." Kenney replied, "That's because I am a computer scientist!" The audience of 120 people laughed a bit, then I answered Kenney's question.

A significant portion of my time at MGH was dedicated to investigating a cohort of 314,292 patients at increased risk for metabolic syndrome. I collaborated with my colleagues at MGH and Harvard to implement a variety of predictive-modeling methods and incorporate text processing techniques to better understand diseases and their complications. We focused on cardiovascular disease, insomnia, and [liver disease](#). This information on this cohort contained the complete clinical details of patients who received care at MGH or Brigham and Women's Hospital (BWH) between 1992 and 2010.

At the end of 2014, cardiologist Dr. Stanley Shaw (who hosted me in his lab during the fellowship), introduced me to Dr. Kathleen Corey, a hepatologist with whom I began a collaboration interrogating the cohort to identify new biomarkers associated with outcomes in individuals suffering from liver diseases and associated comorbidities. Our work yielded several published studies, including one appearing in *The American Journal of Gastroenterology* (see past projects: [researcher.ibm.com/researcher/...rson=ibm-Uri.Kartoun](http://researcher.ibm.com/researcher/...rson=ibm-Uri.Kartoun)).

During this time, I was reading studies by other researchers, all related to liver disease, and trying to identify new research directions. For the first time, I realized the importance of the model of end-stage liver disease

(MELD) risk score, one of the most important and widely used risk prediction scores in medicine. Especially since 2002, MELD has played a crucial role in determining which patient on the waiting list will be next to get a liver transplant.

With the opportunity to access MGH's valuable data and the flexibility of being a fellow, I could explore almost anything. Inspired by the literature I read, I decided to try to identify new biomarkers for cirrhosis, one of the deadliest diseases known to humanity. I thought this was a good research direction given that most previous studies were based on manually selecting a small set of biomarkers to predict mortality.

Inspired by other studies led by Harvard's Informatics for Integrating Biology and the Bedside (i2b2) group, we took an unbiased approach toward discovery of biomarkers. In this approach, a feature-selection machine learning algorithm observed a large collection of health records and identified a small set of variables that could serve as the most efficient predictors for a given medical outcome. We used the well-studied supervised learning paradigm to assess accuracy; we also applied traditional statistical methods to assess the validity of our approach. We realized that by combining the components of MELD (or the components of its extended version MELD-Na) with several easily accessible variables, we could construct a new score that was approximately 10 percent more accurate. We referred to our new score as "MELD-Plus."

Shortly after joining IBM Research, I started working with Kenney, who manages IBM Research's Health Analytics group, to further evaluate MELD-Plus. We deployed the original MGH generalized linear model equation on a database called the IBM Explorys Network. We used a portion of the database representing approximately 18 million patients, pooled from multiple healthcare systems, to further assess the validity of

MELD-Plus.

As we hoped, MELD-Plus proved superior to both MELD and MELD-Na, with an increase in accuracy similar to that found when we applied it to the MGH/BWH database. We believe that our approach of applying machine learning and statistical techniques on large collections of electronic health records to validate it further on an independent data source, may lead to improved care for individuals with liver disease by providing the foundation to devising improved risk scores to better monitor decisions in high-risk patients. We are delighted that *PLOS ONE* has accepted our manuscript for publication, and we hope that our approach may advance the field of hepatology by providing a more accurate tool to assess severity of liver disease.

**More information:** Demonstrating the advantages of applying data mining techniques on time-dependent electronic medical records. American Medical Informatics Association 2015 Annual Symposium, November 14–18, 2015, San Francisco, CA.

Kartoun U. The man who had them all. *ACM Interactions* 2017;24(4):22–3.

Kamath PS, Kim WR. The model for end-stage liver disease (MELD). *Hepatology* 2007;45(3):797–805.

Kartoun U, Corey K, Simon T, Zheng H, Aggarwal R, Ng K, Shaw S. The MELD-Plus: A generalizable prediction risk score in cirrhosis. *PLOS ONE* 2017 (accepted manuscript in production).

Provided by IBM Blog Research

Citation: Machine learning used to surpass MELD in assessing liver disease (2017, October 26) retrieved 11 July 2024 from <https://medicalxpress.com/news/2017-10-machine-surpass-meld-liver-disease.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.