

“Development and validation of a parsimonious AI-based risk score for mortality in heart failure: a UK cohort study” Çalışma Değerlendirmesi

Dr. Hilal Yıldız, Dr. Mehmet Ali Astarcioglu

“Development and validation of a parsimonious AI-based risk score for mortality in heart failure: a UK cohort study” Çalışma Değerlendirmesi

Hazırlayan: Araştırma Görevlisi Dr. Hilal Yıldız, Prof. Dr. Mehmet Ali Astarcioglu
Kütahya Sağlık Bilimleri Üniversitesi Tıp Fakültesi Kardiyoloji Anabilim Dalı

Çalışmanın adı: Development and validation of a parsimonious AI-based risk score for mortality in heart failure: a UK cohort study

Çalışmanın amacı ve metodolojisi: Kalp yetersizliği hastalarında doğru risk stratifikasyonu kritik öneme sahiptir. MAGGIC gibi mevcut basit skorlar sınırlı ayırt edici güce sahiptir (C-index < 0,80) ve ekokardiyografi gibi özel tetkiklere ihtiyaç duymaktadır. Halbuki gelişmiş yapay zeka modelleri daha yüksek doğruluk sunmaktadır. Ancak uygulamadaki zorluklar ve veri erişim gereksinimleri nedeniyle klinik kullanıma entegrasyonları sıklıkla kısıtlanmaktadır.

Kalp yetersizliği hastalarında tüm nedenlere bağlı mortaliteyi öngörmek amacıyla; karmaşık bir yapay zeka modelinden elde edilen içgörüler ve özellik mühendisliği yaklaşımından yararlanarak yalın, yüksek performanslı bir yapay zeka tabanlı risk stratifikasyon geliştirilmesi ve doğrulanması amaçlanmıştır.

Çalışmanın sonuçları ve tartışma: Ahmed N. ve et al. tarafından yapılan çalışmada Clinical Practice Research Datalink (CPRD) Aurum veritabanından elde edilen 373.389 kalp yetersizliği hastasını içeren bir kohortta (≥ 18 yaş; risk skoru geliştirmek için 1.153 İngiliz klinik ve doğrulama için 289 klinik) tüm nedenlere bağlı mortaliteyi öngören bir yapay zeka tabanlı risk modeli geliştirilmiş ve doğrulanmıştır. Risk skorunun performansını karşılaştırmak için, elektronik sağlık kayıtlarına (EHR) uyarlanmış MAGGIC temelli bir Cox orantısız risk modeli referans olarak ele alınmıştır. İlk aşamada, MAGGIC değişkenlerini kullanan ve sağkalım çerçevesinde yapılandırılmış çok katmanlı algılayıcı (MLP) modeli eğitilmiş. Ardından bu model, yüksek prediktif komorbidite değişkenleri ile zenginleştirilmiş ve özellik damıtma (feature distillation) yaklaşımı kullanılarak nihai, optimize edilmiş 11 değişkenli “MLP-Distilled” model elde edilmiştir. Yapay zeka ile geliştirdikleri bu skor; yaş, BKİ, doğum yılı ve maligniteler gibi temel komorbiditeler dahil olmak üzere kolay erişilebilir değişkenleri kullanmaktadır.

MLP-Distilled modelinin, MAGGIC-EHR referans modeliyle karşılaştırıldığında belirgin şekilde daha yüksek ayırt edici performans göstermiş olduğu saptanmıştır. (C-index: 0,801, %95 GA [0,796–0,806] vs. 0,735 [0,729–0,742]). Bu yalın yapay zeka modeli, MAGGIC'e kıyasla daha az sayıda ve daha kolay erişilebilir değişken kullanmaktadır (11 vs. 13 değişken) ve herhangi bir özel tetkik gerektirmemektedir. Model ayrıca kardiyovasküler olaylar ve yeniden hastaneye yatış gibi ikincil sonuçları da doğrulanmıştır.

Ahmed N. ve et al. tarafından yazılan abstractta, yapay zeka algoritması ile kalp yetersizliği hastalarında mortaliteyi öngören hem yalın hem de yüksek performanslı bir risk skoru geliştirilmiştir. Bu yeni yaklaşımın, risk stratifikasyonunun daha iyi yapılabilmesi ve mortalite riski daha yüksek hastaların daha sıkı takibe alınabilmesine yardımcı olması beklenmektedir.

Erişim linki: https://academic.oup.com/ehjdh/article/7/Supplement_1/ztaf143.036/8422927

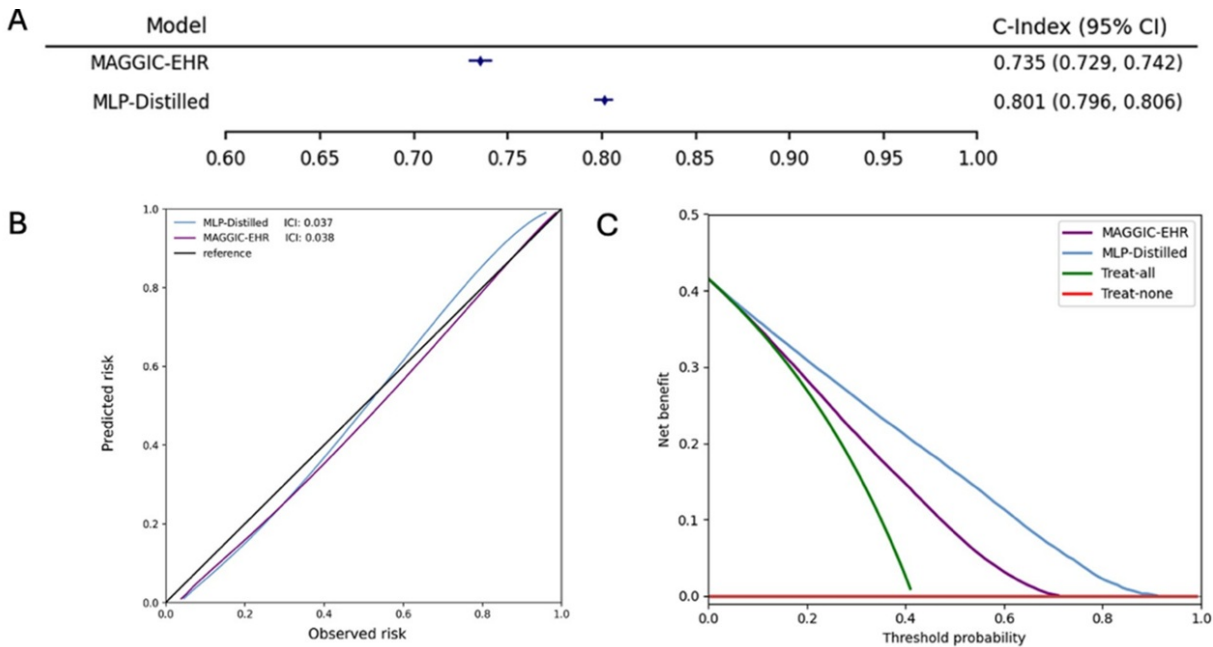


Figure 1: Discrimination, Calibration curves, and net benefit analysis for 12-month risk prediction of all-cause mortality in UK validation data

Discrimination (C-Index) with 95% CI (A), Calibration curves with integrated calibration indices (ICI) (B), and decision curve analysis (C) are presented for all-cause mortality at 12-month timepoint for MAGGIC-EHR and MLP-Distilled Models. For (B) ICI, lower is better with reference (black line) presenting optimal ICI of 0.0.

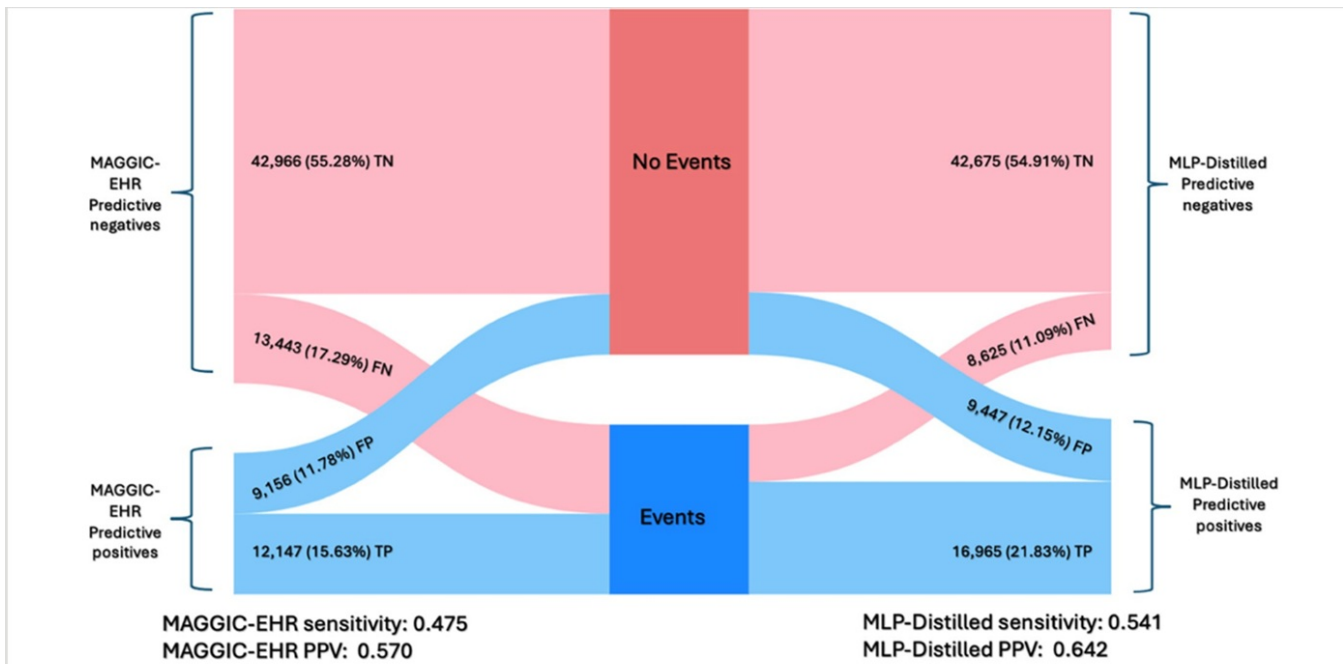


Figure 2. Impact analyses at the 50% decision threshold for 12-month all-cause mortality prediction on UK validation data.

Sankey diagrams compare predicted outcomes between the models, showing how patient classification compares to actual outcomes, "Events" (i.e., death) and "No events" categories at 50% threshold (denoted as dark blue and red respectively). TP: true positive; TN: true negative; FP: false positive; FN: false negatives.