Demographic and socioeconomic factors predict maternal postpartum rehospitalization: a retrospective nuMoM2b dataset study

Colin WAKEFIELD, Martin G FRASCH

Introduction: The advent of artificial intelligence/machine learning (AI/ML) in medicine has opened opportunities for harnessing the power of ML to predict patient outcomes based on diverse data contained in the electronic medical records (EMR). Taken by themselves singular features represented by such diverse data are often not clearly predictive, but combined in an ML modeling framework such so-called weak learners have the power to yield highly predictive models of health outcomes. We hypothesized that sociodemographic and basic, easily obtainable health characteristics of pregnant mothers strongly predict the risk of rehospitalization.

Methods: Using the nuMoM2b dataset from NICHD and reproducible, open access AI/ML techniques we tested and validated this hypothesis.

Results: Among a large (n=10,038) and geographically diverse cohort of nulliparous women with singleton gestations, the present findings show that while some of these characteristics highlight the known disparities in health outcomes (race/ethnicity), most of them are modifiable through social policies and health counseling. These include access to and support during education, comprehensive health insurance, alleviating coming out of poverty, smoking, and unhealthy lifestyle habits.

Conclusion: Across the four ML models deployed, the results are mutually reinforcing yielding a clear cohesive picture. This solution can be used to develop a prospective risk prediction for this important health outcome in pregnant mothers. The strengths of the presented solution are its focus on preventable maternal morbidities and risk factors, demonstration of the impact on adversely affected populations, and reproducible step-by-step executable, documented open-access code with relatively low computational requirements. This enables easy adoption and further development on this and other similar datasets to study the contributions and predictive power of various factors to maternal morbidities.