

Big Data for the Prediction, Prevention and Management of Low Birthweight Babies

David horsburgh

Why:

America is often viewed as a highly developed nation without many of the burdens of developing countries, one of which is infant mortality. However, infant mortality in the US ranks only 55th lowest in the world¹, with a higher number of deaths than countries such as Belarus, Andorra, Cuba and Bosnia and Herzegovina. While this has been declining slowly – there are still challenges due to income diversity and availability of healthcare.

What is not declining is the problem of low birthweight babies, which actually can place an even greater toll on society. Data shows that every 250 grams a baby is under idea weight can cause \$12-16k in first year medical expenses. In contrast, intervening prenatally to result in a normal birth weight could save \$59,700 in first year expenses.

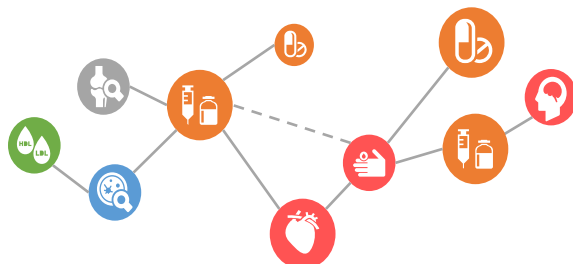
Long term costs beyond this include re-hospitalization and other social service costs – adding to the burden to society and the need for a solution.

Concept and Goal

A solution may be applying a supervised machine learning approach with 2 distinct outcomes

- a. Normal weight at birth (5.5 pounds or greater)
- b. Underweight at birth (sub-5.5 pounds)

This algorithm would look through the available data and identify which are the high-risk indicators, with the eventual goal of creating a proactive model which can provide a real-time risk score for each pregnancy, visible to the healthcare providers managing the mother. This can in turn be use to prioritize prenatal interventions to those at most risk of underweight birth, increasing the efficiency of the healthcare system and reducing societal burden.



Data Required

Data required	Source
Demographic information (age, race, weight)	Medical records
Medical history (smoking, conditions, treatments)	Med records or claims
Familial Medical History (history of underweight)	Familial medical records
Education Level	Education board records
Socioeconomic data	Census, IRS
Social services data (domestic abuse, drug abuse etc)	Social Services / govt. Police Records
Distance from healthcare providers	ISPs, census data, google maps
Environmental data (Water quality, pollution)	Environmental Protection Agency
Behavioural Data (Partying, anti-vaxxer?, etc)	Social Media Platforms
Dietary information	Digital supermarkets, credit card records

Other Uses (Outside of the Predictive / Intervention Model)

Mapping / targeting development plans for community investment

Heat mapping (using google) of home location high risk or previously born underweight babies. This could be used to better target investment in healthcare facilities, access, or education – targeted specifically at high-risk and high-burden communities.



References

1. CIA World Factbook: Infant Mortality Rates (per 1,000 live births). [Cited: March 1st 2019]
Available From: <https://www.cia.gov/library/publications/resources/the-world-factbook/fields/354rank.html>
2. National Healthy Start Association: How Much Does Infant Mortality Cost the Nation? [Cited: March 1st 2019]. Available From: http://www.nationalhealthystart.org/healthy_start_initiative/how_much_does_infant_mortality_cost

