

BAPIOAC19-P05

Machine learning approaches to analyse effect of ethnicity on outcomes following living kidney donation in UK

Shanmugapriya Basker

Cite as: Basker, S. (2020) Machine learning approaches to analyse effect of ethnicity on outcomes following living kidney donation in UK. *The Physician* vol 6; issue 2; epub 01 Jan 2020; DOI: <https://doi.org/10.38192/1.6.2.9>

Article Information

Submitted 01 Nov 2019

Published 01 Jul 2020

Open Access – Creative Commons Licence CC-BY-ND-4.0

Poster presented at BAPIO National Conference, London 2019
Peer reviewed by Subarna Chakravorty Phd & Sunil Daga Phd

Background:

Despite thousands of kidney donations every year across the world, there is decline in donations particularly within the BAME (Black Asian Minority and Ethnic community) and also there is paucity of data for risk assessment within the South Asian ethnicity. This study aims to describe the risk based on ethnicity, gender and regional variations and develop a predictive model.

Method:

This project uses Hospital Episode Statistics (HES) Admitted Patient Care (APC) data from 1997/8 to 2017/18 (20 years) with pseudo-anonymous data filtered by kidney donation codes. After trying several approaches to tackle extremely imbalanced dataset 'Weighted Random Forest' approach was selected as the final model.

Results:

Essential hypertension, End age and deprivation index ranked as three most important variables in risk of

development of hypertension after donation in the Weighted Random Forest. The model has Negative Predictive Value (NPV) of 95% and True Negative Rate (TNR) of 77% proved that it was able to very well predict that the patients had 'No risk of hypertension'. Also the Recall rate of 65% showed the model is not overfitting due to its imbalanced nature. The stability of the model was verified by running the algorithm several times with k-fold cross validation and AUC as well. AUC was between 75% to 76% after running through 100 iterations with different set of data.

Essential hypertension, End age and deprivation index ranked as three most important variables in risk of development of hypertension after donation. Logistic regression analysis showed that pre-existing essential hypertension, male gender, age at kidney donation and low numbers of transplant providers are important risk factors for long term hypertension risk. In sub analysis, we found that South Asian ethnicity and obesity are additional risk factors.