Application number/Title: 41896 - Building Machine Learning Models for Breast Cancer Risk Prediction

Applicant PI: Mr Mahmoud Aldraimli

Application Institution: University of Westminster, London, UK

Keywords provided by the Applicant PI to describe the research project:

cancer, comorbidities, machine-learning, obesity, prediction, risk

Application Lay Summary:

Reducing mortality rates from major NCDs is at the heart of the World Health Organisation (WHO) agenda [1].
The UK NCD profile published by the WHO in 2014 showed that NCDs account for approximately 89% of all mortality [2]. Public Health England published in 2014 that a quarter of the UK population has a long-term medical condition (including major NCDs) and the number of people with multiple conditions is expected to rise [3].
This study aims to identify factors which influence the risk of breast cancer occurrence. This is an important aim as in the long term this could lead to better understanding of the interplay between different illnesses and breast cancer occurrence.
Our approach will be to use Artificial Intelligence (AI) to identify whether there is a link between factors associated with the increased risk of diabetes, obesity and CVD and breast cancer. In particular, we will use a Machine Learning (ML) approach for intelligent data analysis supported by the current digital revolution in collecting and storing data.
Following a systematic review of the UK Biobank, we plan to analyse raw and derived medical and non-medical variables. The analyses will examine whether these variables are correlated with the occurrence of breast cancer, whether these relationships persist in the presence of other variables, and the potential role of obesity, diabetes and CVD in the breast cancer risk prediction.
We will impute missing values while accounting for uncertainty and come up with a predicted risk value of breast cancer. Although the new model will be used to predict breast cancer risk, we will examine our approach for suitability of predicting other NCDs such as obesity, diabetes and CVD. The performance of

each model will be assessed mathematically and clinically.

The results may influence a substantial review of our current public health measures to prevent breast cancer.

The research is funded by the Quintin Hogg Trust, the duration of the research is 36 months as part of a PhD programme.

References:
[1] Global Action Plan for the Prevention and Control of NCDs 2013-2020, available at: http://www.who.int/nmh/events/ncd_action_plan/en/ [Last access: 16/04/2018]
[2] World Health Organization, 2014, Non-communicable Diseases (NCD) Country, Profiles. available at: http://www.who.int/nmh/countries/en/ [Last access: 20/04/2018]
[3] K Fenton, Feb 2014, Health and Wellbeing, Reducing the burden of disease.