

CONFERENCE ABSTRACT

Predictive Models for health services utilization: Current and Future development

17th International Conference on Integrated Care, Dublin, 08-10 May 2017

Stephen Paul Sutch^{1,3}, Chris Morris², Alan Thompson³

1: Johns Hopkins Bloomberg School of Public Health, United States of America;

2: NHS South, Central & West CSU, Oxford, United Kingdom;

3: Johns Hopkins HealthCare, Baltimore, Maryland, United States of America

Introduction: A number of models are available in the US and the UK which predict the risk of hospitalisation, from general and insured populations. These are being used for a variety of purposes including, screening of patients for Case Management Programs, screening for Disease Management Programs, organisational profiling, and assessing financial risk. These uses are in response to health policies to reduce unnecessary hospital admissions, such as of Pay for Performance (P4P) measures, to a genuine need to support populations in avoiding hospital admissions that are both expensive and a patient safety risk.

Method: The predictive models were derived using patient level data, with classification of diagnostic, pharmaceutical and historic utilisation data, using the Johns Hopkins ACG System to reduce the number of variables and provide measures of multimorbidity. The source data includes hospital and primary care service data. Logistic Regressions were undertaken to produce models on the dichotomous outcomes of hospitalisation within 12/6 months, emergency/unplanned hospitalisation within 12 months, rehospitalisation within 30 days, and long stay hospitalisation.

The models were validated using split-half method, and providing ROC analyses to compare different model performance.

Results: A logistic model to predict future hospitalization on UK data produced a C-Statistic of 0.80, which showed an improvement from a age/gender model (0.67) and the existing US model (0.75). For purposes of generating lists of high risk individuals applying a cut-point such that 1% of the population are designated as "positive", the model showed a positive predictive value of 72.77%. An alternative logistic model to predict unplanned hospitalization resulted in a C-Statistic of 0.78.

Discussion: The results demonstrate the importance of casemix classifications to reduce data complexity and provide robust measures of multimorbidity. The models work well in explaining the top 1% and 5% of data, but also perform well in discriminating risk "lower in the population pyramid" to identify potential emerging risk.

Whilst the emphasis of work has been on identifying the highest risk individuals, there is an increased interest in recognising earlier and emerging risk, where more preventative methods can be informed such as chronic disease self-management programs. These models in their current form are being used to identify populations, but work on newly emerging data from Electronic Health Records (EHR), Personal Health Records (PHR), and Social Care data is expected to provide greater insight into these populations and those with highest need. Other models are also focused on other types of utilization across the health care system such as emergency care, outpatient, and primary care visits.

Keywords: utilization risk; predictive modelling; care management
