JOHNS HOPKINS ACG® SYSTEM:

Recalibration of Predictive Models in the ACG System

A Summary of the Recalibration Undertaken in 2020

This short report provides details of the third recalibration of the predictive models within the ACG System. This recalibration was carried out by the Johns Hopkins ACG System Team in collaboration with NHS South, Central & West Commissioning Support Unit (SCW) during the summer of 2020.

The predictive models contained within the Johns Hopkins ACG System were originally derived from United States (US) health care data. In 2006, work at Imperial College and University College London demonstrated that the performance of these models using National Health Service (NHS) data was as robust as the performance in the US.

It is considered good practice to recalibrate predictive models using local data to align them more closely with the way in which data is recorded within that country, how activity is costed and the different ways in which care is provided. For this reason, the ACG System team undertook a major recalibration of the primary predictive models within the ACG System in 2013, resulting in an improvement of the performance of the existing predictive models and the development of a new United Kingdom (UK) model that predicted risk of emergency admission. The exercise was repeated in 2016, again seeing an improvement in the performance of the models.

The ACG System has been in mainstream use in the NHS in England for over 10 years. Currently almost 15 million patients' records are regularly processed through the ACG System software.

ACCEPTED GOOD PRACTICE

There are several approaches to predictive modelling that are accepted as good practice. The table below lists how this exercise and the predictive models within the ACG System conform to these principles.

Practice	Conforms?	Comment
A variety of predictive models is available to allow case finding of different cohorts of people.	Yes	There are more than 12 models that are directly applicable to the NHS.
Models have been calibrated to use UK data, costing and models of care delivery and take into account the ongoing changes to data recording practices.	Yes	
Models recalibrated every 2 – 3 years to keep them fresh.	Yes	Recalibration in 2013, 2016, and 2020.
Recalibration uses a sufficiently large data set that represents the populations to which the predictive models will be applied.	Yes	A representative data set of more than 3 million records was used that combined primary and secondary care data from patients in the SCW geography.
Data used for recalibration and development of new models is split into a 'training' and 'validation' data set to ensure the model is generalisable – i.e. it doesn't just work on the model design's data, it will work on other data sets.	Yes	Models were created using 75% of the data and then tested on the other 25% of the available data.
Results of recalibration and performance of predictive models is published with further details readily available.	Yes	A summary is given in this paper. Further details available upon request.

AIMS & APPROACH

The principal aim of the recalibration exercise was to apply key outputs from the ACG System (version 12.0) as independent variables (predictors) in year 1 to predict individual patient outcomes in year 2. Two main dependent (outcome) variables were used in the study: total cost in year 2, and hospitalisation (inpatient admission) in year 2. The objectives were to:

1. **Recalibrate the existing UK-based predictive models using ACG v12.0 variables**
2. **Compare the performance of the recalibrated models with the existing UK-based models**

The 2020 exercise benefitted from a very large dataset of a little over 3 million records, around 6 times the size of the previous UK recalibration. This brought both benefits, in terms of statistical robustness, but also some disadvantages; the combination of dataset size (3 million rows and nearly one thousand variables) and computational power required in order to enact the analyses, meant that a bespoke virtual machine had to be commissioned for the exercise. Even then, further enhancements were required in order to derive the computational headroom needed. Previous recalibrations had been undertaken in IBM's SPSS software, but following a general move in the NHS towards using the open source 'R' software, the 2020 exercise was undertaken using the latter. Code was developed in R that will enable the recalibration to be undertaken in a semi-automated way next time.

A range of validation analyses were applied in advance of the full recalibration analysis, including visualisation of variables and distributions using Tableau® and Microsoft Power BI, triangulation against reference datasets from prior recalibrations, and joint sense-checking with academic colleagues. There was also a comprehensive set of data preparation exercises undertaken in advance, such as testing cut-points, deriving new variables, creation of dummy variables, and conversion of variable types. The resulting dataset was then also subject to checking that the model build and validation datasets did not differ in any important way.

Two types of regression analyses were used; linear regression for continuous variables such as total cost, and binary logistic regression for yes/no variables such as admission to hospital. For both types of regression, standard statistical measures were applied to confirm that the updated models performed comparably with the existing models, given the same input data.

RESULTS

Unlike other predictive models used in the UK, the ACG System predicts the likely total costs associated with a patient in the coming year in addition to the probability/risk of adverse events such as an emergency admission. The following tables provide a summary of the performance of:

- **The 'linear' models that predict a particular value across a range of values, in this case total cost and pharmacy costs**
- **The 'binary' models that predict the probability of whether a particular event will or will not happen**

Further details of the performance models at various 'cut points' are available upon request.

LINEAR MODELS

Predictive Model	Original US R^2 [1]	UK 2013 R^2	UK 2016 R^2	UK 2020 R^2
Total Cost	0.226	0.256	0.271	0.266
Drug Cost (based on total cost markers)	Not Available	0.355	0.362	0.405
Drug Cost (based on pharmacy cost markers)			0.550	0.598

The performance of predictive models will deteriorate over time due to changes in the way in which health care data is recorded and the introduction of new clinical codes. This recalibration exercise has assured that the total cost model continues to perform well on newer data. There have been further improvements in the performance of the models related to drug costs.

EXISTING BINARY MODELS

Predictive Model	Original US AUC [2]	UK 2013 AUC	UK 2016 AUC	UK 2020 AUC
Any Admission - next 12 months	0.774	0.763	0.780	0.775
Any Admission - next 6 months	0.787	0.782	0.801	0.798
Any Admission – LOS of 12 days or more	Not Available	0.901	0.912	0.903
Emergency (unplanned) Admission	Not Available	0.773	0.786	0.768

The updated models that have been calibrated to work with changes to health data over the last four years continue to perform well. All models are very good at highlighting patients they are designed to identify.

[1] R^2 or R-squared is a measure of forecasting accuracy in linear predictive models. It is the percentage of variation in medical cost explained by the model. A value of 0 indicates that the model explains 0% of the variation, while a value of 1 indicates that the model explains 100% of the variation.

[2] The AUC (Area under the Curve or the Receiver Operating Characteristic [ROC] curve) is one of the performance measures used for binary predictive models. It is the probability that an observation will be classified as a true positive or a true negative, or in other words, correctly. A value of 0.5 suggests the results are no better than chance while a value of 1.0 suggests perfect classification. A value of >0.7 is considered good performance. A value > 0.8 is considered very good.

DISCUSSION & CONCLUSIONS

The regular recalibration of predictive models used in health care is important to account for the ongoing changes to data recording practices.

This is the third time the predictive models in the ACG System have been recalibrated in a collaborative approach between the Johns Hopkins team and the NHS. On this occasion, the data set used to undertake the recalibration comprised six times as many records compared to the previous exercises and the 'R' software was used for the first time. This will make future recalibrations easier to undertake and will support the development of new predictive models.

The recalibrated predictive models will be available to all UK users of the ACG System from January 2021 and will support the work of over 50 Clinical Commissioning Groups (CCGs), several Integrated Care Systems and Sustainability & Transformation Partnerships and numerous Primary Care Networks.

FURTHER INFORMATION

Further information about predictive modelling, the ACG System predictive models and this most recent exercise is available on the Johns Hopkins ACG System website – **www.HopkinsACG.org**.

Acknowledgements

The Johns Hopkins ACG System Team would like to acknowledge and thank NHS South, Central & West CSU for their invaluable support and assistance with this work. Specifically, we would like to thank Chris Morris, Associate Solution Lead and Nessa Becerra Bambury, Senior BI Analyst at SCW for undertaking most of the analytical work.