

Identifying Progressive CKD from Primary Care Records – a Study Protocol

Norman Poh¹, Simon Bull¹, Santosh Tirunagari¹, John Williams², Nicholas Cole², Christopher Farmer³ and Simon de Lusignan³

1: Department of Computer Science, University of Surrey. 2: Department of Clinical and Experimental Medicine, University of Surrey. 3: East Kent Hospitals University NHS Foundation Trust.

Background: One of the key challenges in managing CKD patients is to identify those who are progressive (worsening eGFR) from those who are non-progressive or may even have underlying improvement in their CKD. To this end, we have developed an algorithm capable of identify progressive from non-progressive CKD based on observed historical eGFR trends.

Methods – Clinical codes. We identified 5-Byte Read Codes from the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) for a range of known factors, including codes related to signs and symptoms (*s*), laboratory measurements (*m*) and treatments (*t*) for CKD as well as its associated co-morbidities or risk factors (*r*) such as diabetes, cardiovascular disease and kidney-related diseases. For each risk factor, *s*, *t*, *m* and *r* are systematically identified, leading to a range of plausible phenotype variables for explaining a rapid decline in eGFR

(See our Bayesian Justification accompanying poster). **Algorithm.** We have developed our own regression algorithm – called the broken-stick model – capable of estimating the rate change of eGFR by using a ‘Bayesian’ sliding window of three years in order to provide a stable estimate of the annual rate change of eGFR, while still being sensitive to underlying genuine patterns. The global eGFR slope (annual rate change) of a patient is defined as the average eGFR slope over the patient’s entire history.

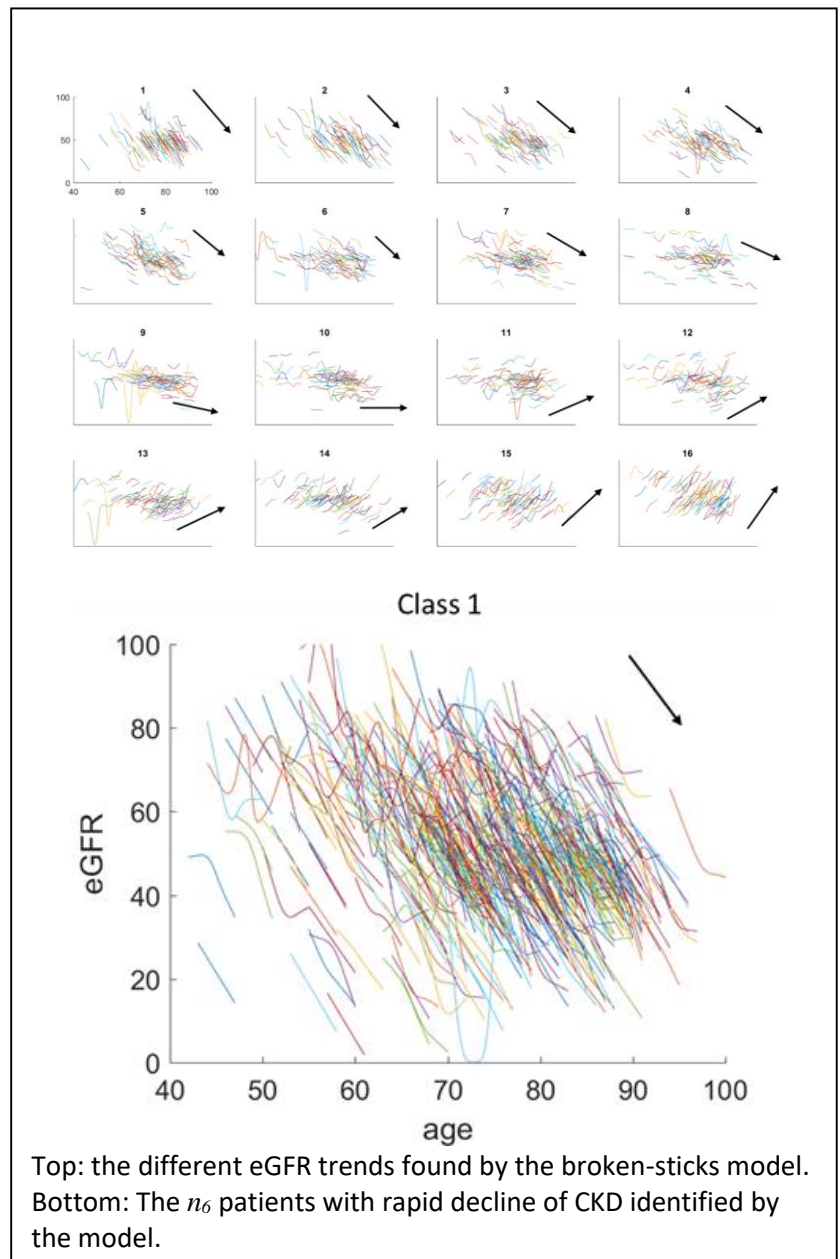
Patient inclusion/exclusion criteria: All patients with eGFR measurements were included. Patients with an acute kidney injury episode or hereditary kidney diseases were excluded. Additionally, patients without *consistent* eGFR trends, defined as having a standard deviation of the eGFR slope of 2 units per year, were excluded.

Initial descriptive observations: The longitudinal observational data was divided into equal groups of approximately 600 patients. In 9 out of 16 groups we found that there was a deterioration in eGFR, one group was equivocal and six groups showed improvement.

Conclusions: A systematic cohort-based retrospective observational study, based on routinely collected primary care data coupled with advanced machine learning algorithms, could improve our understanding of the nature of the rapid progression of CKD in some groups (group 1) of patients in contrast to those (group 16) that show an improvement in eGFR.

Acknowledgement:

This work was supported by the Medical Research Council under the project “Modelling the Progression of CKD”, grant MR/M023281/1. Visit www.modellingCKD.org.



Top: the different eGFR trends found by the broken-sticks model. Bottom: The n_6 patients with rapid decline of CKD identified by the model.