A Factored Co-morbidity Approach for Modelling CKD Progression

Norman Poh[1], Simon Bull[1], Santosh Tirunagari[1], Christopher Farmer[2], Nicholas Cole[3] and Simon de Lusignan[3]

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

Introduction: One of the key challenges in understanding CKD progression is its multifaceted aetiology. This is evident as it is commonly observed that hypertension, heart disease and diabetes are common co-morbidities of CKD. In other words, the existence of co-morbidities can potentially alter the risk of CKD progression, e.g. from stage 3 to 5. Unfortunately, "flat" risk models such as logistic regression, e.g. as implemented by the QKidney score and many similar risk models, are not designed to extract the rich structure induced by a multitude of co-morbidities, the state of which are often captured in routinely collected patient data.

Methods: We revisit the modelling of CKD progression by proposing a Bayesian network that has a certain regular structure, which has attractive properties such as tractability and efficient computation. The risk of progression to more severe CKD stages can be inferred not only from eGFR but signs and symptoms (s) related to kidney impairment, while causality of CKD is captured by risk factors (r) and treatment strategies (t). Each co-morbidity is then allowed to directly interact with the CKD progression risk only if the diagnosis of the co-morbidity is available. Where this information is not available, the state of the co-morbidity can be inferred from its own s, r, and t variables.

In addition, the model can also incorporate past episodes of acute kidney injury (AKI) when estimating the risk of CKD progression. This exploits the fact that people with AKI are more likely to have a lower eGFR value. The extent of this degradation is explicitly modelled by the Bayesian network. When the AKI information is not available, the risk of AKI can be inferred from its own r, s, and t variables, in a process known as 'marginalisation'. Therefore, the interplay between CKD and AKI can be captured by the model.
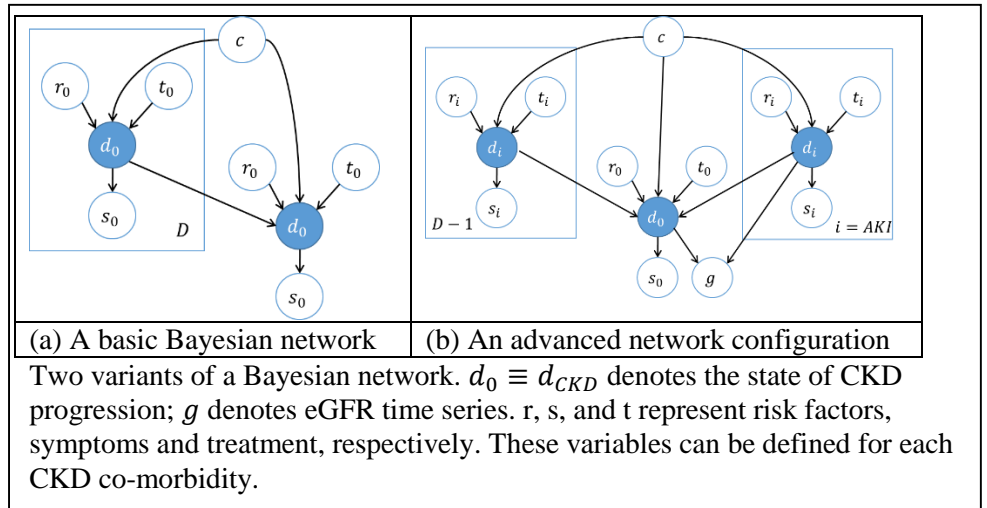
Analysis: As a proof of concept, the proposed Bayesian network can be verified "by parts". In essence, the factored co-morbid approach offers a divide-and-conquer strategy to implement risk models. Some of the relationships of the model are informed by the literature and can be expressed in the conditional probabilities below.

Statements of fact	Possible local models
eGFR is influenced by the presence of AKI episodes	$p(g\|d_{AKI})$
CKD progression is reflected by eGFR	$p(g\|d_{CKD})$
Heart disease, hypertension and diabetes are co-morbidities of CKD	$p(d_{CKD}\|d_{CVD}, d_{HT}, d_{diabetes})$
Heart disease depends on vascular risk factors or measurable signs, is influenced by treatment and can be assessed by symptoms.	$P(d_{CVD}\|r_{CVD}, s_{CVD}, t_{CVD},)$

The local models can be linked to form a single Bayesian network as shown in figure (a) below. In (b), the model is tailored for AKI, which makes provision for eGFR (g) to be influenced by AKI directly.

Conclusion and relevance: Current risk models are 'flat' and do not exploit the rich information captured in patient records. The proposed factored co-morbidity approach offers a natural next step whilst being tractable and efficient in computation. The merit of this model will be investigated as part of the MRC Modelling CKD project using data extracted from the Royal College of GPs database.

(a) A basic Bayesian network	(b) An advanced network configuration

Two variants of a Bayesian network. $d_0 \equiv d_{CKD}$ denotes the state of CKD progression; g denotes eGFR time series. r, s, and t represent risk factors, symptoms and treatment, respectively. These variables can be defined for each CKD co-morbidity.