AJKD risk score editorial

Celine Foote[1,2] MD PhD, Mark Woodward[1,3,4] PhD, Meg J Jardine[1,2] MD PhD

1The George Institute for Global Health, University of Sydney, Australia

2Concord Repatriation General Hospital

3The George Institute for Global Health, University of Oxford, UK

4Department of Epidemiology, Johns Hopkins University, Baltimore MD, USA

Corresponding author:

Meg J Jardine.

The George Institute for Global Health

Lvl 10, KGV Building, Missenden Road, Camperdown NSW 2050 Australia

mjardine@georgeinstitute.org.au

Title

Scoring Risk Scores: Considerations before incorporating clinical risk prediction tools into your practice

Dialysis units are increasingly grappling with considerations of the likely outcomes of dialysis therapy in older people. Patients aged 75 or more are the fastest growing age group in dialysis internationally with an overall increase of 57% in the US over the last decade[1] _ENREF_1. More than one fifth of incident dialysis patients in Australia and the US[1,2] are aged 75 or over. Moreover, outcomes for older people commencing dialysis are poor. _ENREF_4Mortality rates for older patients (≥65) receiving dialysis in the US are twice that of people with diabetes, cardiovascular disease, heart failure or cancer[1]. The actuarial life expectancy of a 75 year old

Australian on dialysis is only one third that of population-based controls[3] _ENREF_5. Survival varies substantially among older dialysis patients, at least in part due to the considerable heterogeneity in comorbidity and geriatric syndromes[4]. There is increasing uncertainty over whether dialysis therapy improves survival in all older patients, particularly those with multiple co-morbid conditions. Beyond survival, dialysis therapy imposes a time and lifestyle burden and has been associated with higher rates of functional[5] and cognitive[6] loss and decline in quality of life[7] in older patients.

These considerations can make it difficult to advise older patients on the potential benefits and harms of dialysis therapy. Risk prediction models (or 'risk scores') are designed to predict an individual's probability of an adverse outcome, such as premature death, taking account of relevant demographic, clinical and other information.._ENREF_8 They may assist clinicians and patients in decision-making[8].

In this issue of AJKD, Wick et al make a valuable contribution to our understanding of factors underlying short-term mortality in older dialysis patients. They examine 6-month survival outcomes in a cohort of 2199 adults aged ≥ 65 initiating dialysis in Alberta, Canada between May 2003 and March 2012 defined by administrative and clinical data. Within this cohort, they identify independent predictors of mortality which include age ≥80 years, 'early' dialysis initiation (defined by estimated glomerular filtration rate at dialysis initiation), atrial fibrillation, congestive cardiac failure, lymphoma, metastatic cancer and hospitalisation in the prior 6 months. From this they have developed a risk score that potentially could be used to estimate mortality risk over the next 6 months for an individual older patient commencing dialysis.

The Alberta risk score adds to existing risk scores for short-term mortality prediction in people commencing dialysis. Other mortality risk scores for older patients include two French risk score derived from the national renal registry (REIN),[9,10] _ENREF_9 each identifying nine independent predictors, and the US risk score derived from Medicare and Medicaid data, identifying seven independent predictors[11]. A limitation to all the existing dialysismortality scores is their inherent selection bias. They are all derived from patient populations who have commenced dialysis and do not include those who decline, are not selected for, or do not survive to, dialysis initiation. A score that evaluates older patients at the point of decision making, rather than at the point of starting dialysis, would be very valuable.

Before adopting a risk score into practice, clinicians need to decide whether it will assist predictions for the outcomes of their patients or merely provide an insight into the setting in which it was derived. During these considerations, clinicians may find it helpful to pose three questions.

Did the risk score predict well in the study population in which it was developed?

Predictive performance is traditionally assessed by determining the score's 'discrimination' and 'calibration'[8,12] _ENREF_14. *Discrimination* is a measure of ranking - the score's ability to order patients correctly in terms of their risk - whilst *calibration* is a measure of scaling – how close the distribution of predicted risks from the score is to the true distribution of risks. The most familiar measure of discrimination is the concordance statistic (*c-statistic*) derived from a logistic regression model - a measure of correlation that is equal to the area under the receiver-operating characteristic curve[12,13]. The c-statistic ranges from 0.5 (no better than chance) to a theoretical maximum of 1 (perfect concordance). In vascular research, c-statistics of

0.7 and 0.8 are commonly considered to represent ‘reasonable’ and ‘strong’ discrimination, respectively[14]. The c-statistics of the Alberta, French and US risk scores all discriminate reasonably well at 0.72, 0.75 and 0.72, respectively. Discrimination is not the same as statistical significance[15]. For example, our model, developed from the ANZDATA cohort (an older Australian dialysis population), identified a number of independent patient and practice predictors of survival that were clearly statistically significant but had a c-statistic of only 0.614 indicating that much of the variation in outcome was determined by unmeasured factors[16].

Calibration, the agreement of observed and predicted outcomes, is most appropriately assessed using a calibration plot[12], which plots the observed and expected (according to the risk score) frequencies of events for a set of mutually exclusive subgroups, typically defined by deciles, of the expected frequencies. The points will describe a diagonal straight line if agreement is good.

Does the score accurately predict outcomes in people like my patients?

Most risk scores will have been internally validated, by using methods such as cross-validation - splitting the original cohort into two (or more) subpopulations, developing the score in one group and testing in the other- or bootstrapping[12]. However, the clinician will want to know the generalisability of the risk score; that is, how well it performs in other populations (e*xternal validity),* most specifically the population they are treating. The Alberta score is yet to be replicated in a truly independent population. The question of external validity is not just one to concern methodologists. An illustrative example can be taken from the $CHADS_2$ score, a risk score commonly used for assessing the risk of stroke in non-valvular atrial fibrillation (AF)[17]. This score was derived from a pooled set of studies, predominantly of cardiology patients, and did not consider kidney function. As a consequence, the

highest achievable $CHADS_2$ score (6) corresponds to an annual stroke rate of 18.2%, which is likely to be less than the average risk for most dialysis patients with AF[18]. Therefore, use of the $CHADS_2$ score among such patients would likely lead to a considerable underestimate of risk.

In the absence of *validation* studies in external dialysis populations, clinicians caring for dialysis patients should consider how their own patient population and setting compares to the population(s) wherein the score was developed. Overall mortality rates are a simple means of comparison. The Alberta cohort has a 6 month mortality rate of 17% compared with 19% in the French cohort[10]. The US cohort has a 3 month mortality rate of 12%[11] while the ANZDATA cohort has a 12-month mortality rate of 23%[16], Baseline characteristics should also be used for comparison purposes. All three scores include a measure of cardiovascular disease. Features of the Alberta cohort that may not be typical of all dialysis programs are the presence at dialysis initiation of lymphoma in 4% of participants and of metastatic malignancy in 3.2% while a large minority (39.5%) commenced dialysis with an eGFR of 10 ml/min/1.73 m^2 or greater.

Will using the score make a meaningful contribution to decision making for me or my patients?

Clinical utility reflects usability and the extent to which the risk score actually affects clinical decisions. Most risk scores are opportunistically derived from large datasets developed for other purposes and so may not be able to assess all potentially important factors. Risk scores from the general population may have good predictive power but be so dominated by age that they represent low clinical utility in decision making for, and with, an individual older patient. In part, clinicians can assess the likely clinical utility by looking at the extent to which the source dataset was able to

test the factors that are important in the clinician's own setting. All the scores incorporate some measure of cardiovascular disease, reflecting how common these conditions are among older dialysis patients. The Alberta score incorporates the presence of lymphoma and metastatic malignancy. It will have most clinical utility in settings where these conditions appear with reasonable frequency. The French and US scores incorporate measures of functional status, dependency and requirement for supported residence_ENREF_18. Other considerations are ease of use. The 'surprise' question is a predictive variable that is very simple to administer[19] while the Alberta score includes measures of health service utilisation.

With increasing numbers of older adults developing chronic kidney failure, an improved understanding of the factors that underpin outcomes following initiation of dialysis is imperative. The Alberta 6-month mortality risk score adds to the current set of scores by producing a rigorously derived model, providing further insight into important variables and potentially informing treatment decisions of patients and their clinicians. However, it is likely that no single risk score will be suitable for the multitude of diverse dialysis settings given their inherent reflection of their original setting. Large dialysis providers contemplating incorporating risk scores into practice should consider conducting comparative studies of the predictive power and clinical utility of different scores. Ideally, prospective studies would be undertaken testing the impact of potentially important factors on life expectancy and the patient experience in older patients with chronic kidney failure adults, both those initiating dialysis and those choosing a supportive (non-dialysis) pathway. Such a study would greatly facilitate meaningful discussions of the full impact of treatment choice on older patients.

Funding:

CF is supported by… MW is supported by …..

MJJ is supported by a Career Development Fellowship from the National Health and Medical Research Council of Australia and the National Heart Foundation.

1. USRDS 2013 Annual Data Report: Atlas of End-Stage Renal Disease in the United States. Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2013.
2. ANZDATA. Registry Report Volume 34. Adelaide, South Australia: Australia and New Zealand Dialysis and Transplant Registry; 2011.
3. Foote C, Ninomiya T, Gallagher M, et al. Survival of elderly dialysis patients is predicted by both patient and practice characteristics. Nephrol Dial Transplant 2012;27:3581-7.
4. Bowling CB, O'Hare AM. Managing older adults with CKD: Individualized versus disease-based approaches. American Journal of Kidney Diseases 2012;59:293-302.
5. Jassal SV, Chiu E, Hladunewich M. Loss of independence in patients starting dialysis at 80 years of age or older. New England Journal of Medicine 2009;361:1612-3.
6. Murray AM, Tupper DE, Knopman DS, et al. Cognitive impairment in hemodialysis patients is common. Neurology 2006;67:216-23.
7. Da Silva-Gane M, Wellsted D, Greenshields H, Norton S, Chandna SM, Farrington K. Quality of life and survival in patients with advanced kidney failure managed conservatively or by dialysis. Clinical Journal of The American Society of Nephrology: CJASN 2012;7:2002-9.
8. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: A framework for traditional and novel measures. Epidemiology 2010;21:128-38.
9. Couchoud CG, Beuscart JBR, Aldigier JC, Brunet PJ, Moranne OP. Development of a risk stratification algorithm to improve patient-centered care and decision making for incident elderly patients with end-stage renal disease. Kidney International 2015;88:1178-86.
10. Couchoud C, Labeeuw M, Moranne O, et al. A clinical score to predict 6-month prognosis in elderly patients starting dialysis for end-stage renal disease. Nephrology Dialysis Transplantation 2009;24:1553-61.
11. Thamer M, Kaufman JS, Zhang Y, Zhang Q, Cotter DJ, Bang H. Predicting early death among elderly dialysis patients: Development and validation of a risk score to assist shared decision making for dialysis initiation. American Journal of Kidney Diseases 2015;66:1024-32.
12. Woodward M. Study Design and Data Analysis. Third ed: CRC Press, Boca Raton; 2014.
13. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. BMC Med Res Methodol 2012;12:82.
14. Hosmer D, Lemeshow S. Applied Logistic Regression (2nd Edition). New York: John Wiley & Sons; 2000.
15. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928-35.
16. Foote C, Ninomiya T, Gallagher M, et al. Survival of elderly dialysis patients is predicted by both patient and practice characteristics. Nephrology Dialysis Transplantation 2012.
17. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: Results from the National Registry of Atrial Fibrillation. Journal of the American Medical Association 2001;285:2864-70.

18. Reinecke H, Brand E, Mesters R, et al. Dilemmas in the management of atrial fibrillation in chronic kidney disease. Journal of the American Society of Nephrology 2009;20:705-11.
19. Cohen LM, Ruthazer R, Moss AH, Germain MJ. Predicting six-month mortality for patients who are on maintenance hemodialysis. Clinical Journal of the American Society of Nephrology 2010;5:72-9.