

Frailty and mortality: 'Same-same but Different'

Sei J Lee^{1,2}

¹San Francisco VA Quality Scholars, San Francisco, California, USA

²Division of Geriatrics, University of California, San Francisco (UCSF), San Francisco, California, USA

Correspondence to

Dr Sei J Lee, San Francisco VA Quality Scholars, San Francisco, CA 94121, USA; sei.lee@ucsf.edu

Accepted 7 December 2018

Prediction models hold tremendous promise as a way to improve patient outcomes and healthcare quality more generally by efficiently targeting interventions to those patients most likely to benefit. Most aspects of healthcare (ie, tests, medications and procedures) have associated risks and burdens. Accurate prediction of patients at higher risk for adverse outcomes from their underlying conditions allows for better targeting of interventions, so that only patients for whom the benefits of the intervention outweigh the risks of treatment are provided the intervention. Conversely, accurate prediction of patients at low risk for adverse outcomes (for whom the risks of the intervention outweigh the benefits) can help those patients avoid unnecessary and potentially harmful interventions. Thus, accurate prediction models play a pivotal role in the vision of personalised medicine, where all clinical decisions are tailored to each individual's unique risk profile.¹

In this issue of *BMJ Quality & Safety*, McAlister and van Walraven² expand our knowledge of prediction for common, adverse hospitalisation outcomes (prolonged hospitalisation, 30-day mortality and readmission) in older adults. They use province-wide data from Ontario, Canada in 2004–2010 to compare how two previously published prediction models (Hospital Frailty Risk Score or HFRS³ and Hospital-patient One-year Mortality Risk or HOMR⁴) predict these outcomes in historical Ontario data. They found that the HFRS more accurately predicted prolonged hospitalisation, while the HOMR more accurately predicted 30-day mortality. Both HFRS and HOMR poorly predicted 30-day readmissions to hospital.

The authors should be commended for conducting a methodologically rigorous external validation of previously

developed prediction models. Previous reviews have found that many prediction models are developed, but few are externally validated.⁵ Without external validation, providers face substantial uncertainty about whether a model is accurate enough to guide clinical decisions. Thus, methodologically rigorous validation studies such as this one constitute a critical but often-ignored component in the chain of evidence that starts with prediction model development and ends with prediction models being used to inform clinical care.

As the authors themselves note, their primary finding, that the HFRS better predicts prolonged hospitalisation while the HOMR better predicts 30-day mortality, comes as no surprise. The HFRS was developed and optimised to predict frailty and prolonged hospitalisation. In contrast, the HOMR was developed and optimised to predict 1-year mortality. This study's results show that while adverse outcomes such as prolonged hospitalisation and mortality often cluster together, they are distinct: A prediction model optimised for 1-year mortality predicts 30-day mortality better than a prediction model optimised to predict frailty and prolonged hospitalisation. An optimistic interpretation of these results is that prediction models have reached a level of sophistication where related outcomes such as mortality and prolonged hospitalisation can be distinguished and distinct prediction models are needed for these related (same-same) but distinct (different) outcomes.

Three additional factors should be considered when interpreting the results of this study²: (1) differences between the UK and Ontario, (2) the importance of calibration as well as discrimination in the validation of prediction studies, and (3) the importance of physical function



► <https://doi.org/10.1136/bmjqs-2018-008661>



© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Lee SJ. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2018-008821

in the prediction of outcomes for hospitalised older adults.

UK/ONTARIO DIFFERENCES

One striking result of this study concerns the differences between the UK and Ontario populations of hospitalised older adults. At baseline, hospitalisation occurred much more commonly in the UK, with 40.1% of UK patients experiencing multiple prior admissions in the previous 2 years compared with only 6.3% in Ontario. This profound difference in the hospitalisation exposure likely led to increased rates of the International Classification of Diseases, 10th Revision (ICD-10) codes in the UK, resulting in the UK cohort having higher Charlson comorbidity scores (2.9 in UK, 2.0 in Ontario) and higher proportion of frailty (58% in UK, 26% in Ontario). These baseline differences in hospitalisation exposures may also have contributed to some of the unexpected, surprising results of this study. For example, patients with higher frailty risk scores (HFRS) in the UK were *more* likely to be readmitted; however, patients in Ontario with higher frailty risk scores were *less* likely to be readmitted. Future studies should explore potential sources of the differences in hospitalisation rates in the UK and Ontario, including the differential use of geriatric day hospitals or different coding practices. This information could be very helpful in interpreting some of these intriguing differences in both the rate of hospitalisation and the apparent differential effect of frailty on subsequent readmissions.

CALIBRATION AS WELL AS DISCRIMINATION

Future studies should focus on both discrimination and calibration as coequal components in prediction model validation.⁶ Historically, discrimination (usually measured by the c-statistic) has overshadowed calibration in prediction model validation. This is unfortunate. Calibration is as important as discrimination (if not more important) for prediction models that are going to be used to inform clinical decisions.^{7 8} Discrimination measures how well a model stratifies or orders patients by risk. However, patients and providers are less interested in whether a patient is higher (or lower) risk than others. The information that is most helpful in clinical decision making is absolute predicted risk, which is evaluated through calibration. For example, knowing that a patient is in the highest quintile of risk is often less important than knowing that the patient's 1-year risk of the outcome is 50%. Thus, prediction model validation studies should prominently display calibration (predicted vs observed outcome rates) across risk groups so that readers can evaluate model calibration and calculate the predicted risk for an individual patient with a specific set of predictors.

PHYSICAL FUNCTION

There is an extensive literature highlighting the importance of physical function as a predictor of outcomes for hospitalised older adults.⁹ However, neither the HFRS or HOMR considers functional status predictors. This omission likely reflects the fact that functional data are often unavailable in electronic and administrative databases. Thus, although the HFRS, the HOMR and the current validation study are methodologically sound, major advances in the prediction of hospitalisation outcomes for older adults will likely remain out of reach until we can access additional data on factors such as physical function. Given the intrinsic importance of physical function to patients and its additional value in predicting outcomes such as readmission and nursing home placement, I urge regional and national data systems to follow the lead of the US Department of Veterans Affairs in routinely collecting functional data.¹⁰

CONCLUSION

This is an exciting time for clinical prediction with several trends coalescing to make prediction easier, faster and more accurate. First, the increasingly widespread use of electronic medical records means that more and more data are becoming readily accessible for clinical prediction. Instead of relying solely on administrative data such as age and ICD-10 codes, clinical data such as laboratory results, radiology results and pharmacy data are increasingly being used to identify which patients are at highest risk.¹¹ Second, there is increasing acceptance by clinicians and the public that 'big data' and 'predictive analytics' can make many things, including healthcare, better. While previous generations of clinicians viewed sophisticated 'black-box' prediction models with scepticism, newer generations of clinicians, growing up in an age where Google correctly guesses your search phrase after three letters and Netflix recommends a show that ends up being your favourite, are more comfortable using predictions from sophisticated models to inform clinical decisions. These trends suggest that clinical prediction models will play a larger role in healthcare in the future. Studies such as this one will be a critical component of the evidence base that ensures that clinical prediction models fulfil their promise of better, safer care.

Funding SJL was supported by the National Institute on Aging (R01AG047897 and R01AG057751), VA HSR&D (IIR 15-434) and VA Office of Academic Affiliations (VA Quality Scholars Program, Grant #AF-3Q-09-2019-C).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES

- Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793–5.

- 2 McAlister F, van Walraven C. External validation of the Hospital Frailty Risk Score and comparison with the Hospital-patient One-year Mortality Risk Score to predict outcomes in elderly hospitalised patients: a retrospective cohort study. *BMJ Qual Saf* 2018. doi: 10.1136/bmjqs-2018-008661. [Epub ahead of print 31 Oct 2018].
- 3 Gilbert T, Neuburger J, Kraindler J, *et al.* Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. *Lancet* 2018;391:1775–82.
- 4 van Walraven C. The Hospital-patient One-year Mortality Risk score accurately predicted long-term death risk in hospitalized patients. *J Clin Epidemiol* 2014;67:1025–34.
- 5 Yourman LC, Lee SJ, Schonberg MA, *et al.* Prognostic indices for older adults: a systematic review. *JAMA* 2012;307:182–92.
- 6 Alba AC, Agoritsas T, Walsh M, *et al.* Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA* 2017;318:1377–84.
- 7 Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem* 2008;54:17–23.
- 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 Covinsky KE, Pierluissi E, Johnston CB. Hospitalization-associated disability: "She was probably able to ambulate, but I'm not sure". *JAMA* 2011;306:1782–93.
- 10 Brown RT, Komaiko KD, Shi Y, *et al.* Bringing functional status into a big data world: Validation of national Veterans Affairs functional status data. *PLoS One* 2017;12:e0178726.
- 11 Wang L, Porter B, Maynard C, *et al.* Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. *Med Care* 2013;51:368–73.