Title: **Predicting 1-year mortality in older hospitalized patients: external validation of the HOMR model**

Running title: **External validation of the HOMR model**

- Denis Curtin MB[1,2]
- Darren L Dahly , PhD[3]
- Maarten van Smeden, PhD[4]
- Desmond P O’ Donnell, MB[2]
- David Doyle[5]
- Paul Gallagher, PhD[1,2]
- Denis O’ Mahony, MD[1,2]

1. Department of Medicine, University College Cork, Cork, Ireland.
2. Department of Geriatric Medicine, Cork University Hospital, Wilton, Cork, Ireland.
3. HRB Clinical Research Facility Cork, School of Public Health, University College Cork, Ireland.
4. Department of Clinical Epidemiology, Leiden University Medical Centre, Netherlands.
5. Department of Information and Communications Technology, Cork University Hospital, Wilton, Cork, Ireland.

Corresponding author:

Dr. Denis Curtin

Department of Geriatric Medicine, Cork University Hospital, Cork, Ireland, T12 DC4A

Email: DenisCurtin@physicians.ie

Word count:

Abstract: 298

Main text: 1992

Abstract

Background

Accurate prognostic information can enable patients and physicians to make better healthcare decisions. The Hospital-patient One-year Mortality Risk (HOMR) model accurately predicted mortality risk (concordance [c] statistic 0.92) in adult hospitalized patients in a recent study in North America. We evaluated the performance of the HOMR model in a population of older inpatients in a large teaching hospital in Ireland.

Design

Retrospective cohort study.

Setting

Acute hospital

Participants

Patients aged ≥65 years cared for by inpatient geriatric medicine services from January 1st 2013 to March 6th 2015 (n = 1654). After excluding those who died during the index hospitalization (n = 206), and those with missing data (n = 39), the analytical sample included 1409 patients.

Measurements

Administrative data and information abstracted from hospital discharge reports were used to determine covariate values for each patient. One-year mortality was determined from the hospital information system, local registries, or by contacting the patient's general practitioner. The linear predictor for each patient was calculated and performance of the model was evaluated in terms of its overall performance, discrimination, and calibration. Recalibrated and revised models were also estimated and evaluated.

Results

One-year mortality rate after hospital discharge in this patient cohort was 18.6%. The unadjusted HOMR model had good discrimination (c statistic 0.78; 95% confidence interval [CI] 0.76 -0.81) but was poorly calibrated and consistently overestimated mortality prediction. The model's performance was modestly improved by recalibration and revision (optimism corrected c-statistic 0.8).

Conclusions

The superior discriminative performance of the HOMR model reported previously was substantially attenuated in its application to our cohort of older hospitalized patients, who represent a specific subset of the original derivation cohort. Updating methods improved its performance in our cohort, but further validation, refinement and clinical impact studies are required prior to use in routine clinical practice.

Introduction:

An important principle when caring for an older person with frailty and multi-morbidity is to align interventions to the patient's condition, preferences, and prognosis.[1] When life expectancy is limited, strategies to optimize quality of life may be prioritized over invasive or futile interventions. Discussions about goals of care, however, are often deferred in frailer older patients because of the uncertainty associated with prognostic estimates.[2] An accurate method of assessing prognosis could inform and motivate discussions between physicians and their patients about values, priorities, and therapeutic goals.

The Hospital-patient One-year Mortality Risk (HOMR) model has been shown recently to accurately predict one-year mortality risk in hospitalized patients.[3, 4] It is comprised of covariates that include demographics, co-morbidities, severity of acute illness, and recent acute hospital care utilization (**Supplementary Appendix S1**). These covariates are determined at the time of hospital admission using routinely collected health administrative data. Over 3 million patients aged 18 or older were included in the validation studies in Ontario and Alberta (Canada), and Boston (United States).[3, 4] The HOMR model had a very high discriminative performance (concordance [c] statistic of 0.89 -0.92) and there was a less than 1% difference between the observed and expected percentages of deceased patients at 1 year.

To our knowledge, the HOMR model's performance exceeds that of other similar prognostic models. However, it has not been validated in an exclusively older (≥65 years) hospitalized patient population. The aim of this study was to evaluate the performance of the HOMR model in a population of older hospitalized patients in a large teaching hospital in Ireland.

Methods:

Data collection

The HOMR model was retrospectively applied to all hospitalized patients aged 65 years or older that were under the care of the specialist geriatric medicine service in Cork University Hospital from January 1st 2013 to March 6th 2015. When patients were admitted more than once during that period, a single hospital admission was chosen at random as the index hospitalization. Most of the information required to calculate the HOMR model was obtained using administrative data from the Hospital In-Patient Enquiry system (HIPE -a national database of coded discharge summaries). The *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification* (ICD-10-AM), Australian Classification of Health Interventions (ACHI) and *Australian Coding Standards* (ACS) apply to all activity coded in HIPE in Ireland.[5] Details about home supports prior to admission as well as provision of home oxygen therapy, which are not routinely collected by administration staff in Ireland, were obtained from the consultant geriatrician discharge reports. When information was missing from these sources, the patients' medical records were reviewed. Covariate values were determined independently by two researchers with discrepancies resolved through consensus.

Deaths within one year of hospital admission were determined by accessing the hospital clinical information system, an online death notification system (https://www.RIP.ie), the Births, Deaths and Marriages Registry Office in Cork City, and, if required, by contacting the patient's general practitioner. Unlike the initial HOMR derivation and validation studies, patients who died during the index hospital admission were not included. There were two reasons for this. Firstly, geriatrician discharge reports were used to obtain information about home supports for the HOMR model, and these details were generally not included when the patient died during hospitalization. Secondly, the value of the predictive model, for the present project, is to calculate 1-year mortality risk after the acute hospital episode. Predicting in-hospital deaths largely depends on specific clinical factors.

Statistical analysis

A sample size that results in at least 100 events, and preferably 200 or more events, is recommended to externally validate a prognostic model.[6] We estimated that one-year mortality *after* hospital discharge would very likely exceed 15%,[7, 8] and on that basis calculated that a sample size of 1400 patients would be required.

To validate the HOMR model, the linear predictor for each patient was calculated based on the coefficient values provided in Appendix E of the original HOMR model development study.[3] The HOMR model was then evaluated in terms of its overall performance, discrimination and calibration. The model's overall performance was evaluated using the Brier score, rescaled to range from 0 to 1, with higher values indicating better performance.[9] Discrimination, which refers to how well the model distinguishes those with the outcome from those without the outcome (i.e. death in this case), was measured using the c statistic. Calibration refers to the agreement between observed outcomes and predicted outcomes and is usually displayed using a calibration plot. In addition to calibration plots, we also report the maximum and average difference in predicted versus loess-calibrated probabilities (Emax and Eavg).[10] Finally, we report bootstrapped 95% confidence intervals for these metrics, based on 500 resampled replicates.[11]

To recalibrate the HOMR Model, three additional logistic regression models were estimated.[12] The first additional model included the HOMR linear predictor, with its coefficient set to equal 1, and a freely estimated intercept (**Recalibration in the Large**). The second model then allowed the coefficient on the HOMR linear predictor to be freely estimated (**Logistic Recalibration**). The third model included the complete set of variables used in the HOMR model, including the same transformations and interactions, and allowed their respective coefficients to be freely estimated (**Model Revision**). The performance of each of these models was assessed using the same metrics used to validate the original HOMR model. In addition, optimism corrected c-statistic and shrinkage factor were estimated for the Model Revision using bootstrapping (with 500 re-sampled replicates).

All analyses were conducted using R language for statistical computing software, [13] version 3.4.3 (2017-11-30). All data and the code used to analyze it and generate outputs can be found on the Open Science Framework (https://osf.io/tv26k/).

Results:

Baseline characteristics of study population

Between January 1st 2013 and March 6th 2015, 1654 individual patients aged 65 year or older were hospitalized under the care of the specialist geriatric service. Of these, 206 patients (12.4%) died during the index hospitalization and therefore were not included in the analysis. After removing 39 patients with missing outcome data (2.7%), a final sample of 1409 patients was analysed. Of these, 259 (18.4%) died within 1 year of admission to hospital. The median age of the study patients was 80 years (interquartile range 74 -85), two thirds were living independently prior to their hospital admission, and 94.5% were admitted through the emergency department. The baseline characteristics of the study participants are summarized in **Table 1**.

HOMR model external validation

When the HOMR model was applied directly to the sample of 1409 older patients, it showed good discrimination (c statistic =0.78). Calibration, however, was poor (see **Figure 1** for calibration plot) with the model consistently over-estimating mortality at all but the lowest levels of risk (see **Table 2** for performance metrics).

Performance of updated HOMR model

All three updating methods improved calibration over the original model. Recalibration in the Large resulted in a lower intercept (-0.42; see **Table 2**) and a significant improvement in model fit over the HOMR model (likelihood ratio test [LRT] Chi-square p value= <0.001). Logistic Recalibration did not lead to additional improvements in model fit (LRT Chi-square p value = 0.85), with a recalibration slope of 0.99 (i.e. close to 1). The Brier score and Eavg were improved by recalibration (**Table 2**). The calibration plot for Recalibration in the Large (which is virtually identical to the plot for Logistic Recalibration) is shown in **Figure 1**. In addition to improving calibration, Model Revision also improved discrimination (c statistic =0.82). The optimism corrected c-statistic for the Model Revision was 0.8, and the shrinkage factor was 0.91, indicating some overfit. The re-estimated HOMR model, with regression coefficients, is shown in **Supplementary Appendix S2**.

Discussion:

This study provides information about the performance of the HOMR model in new patients, in a different geographical region, when validated by investigators who were not involved in the model's development. The high discriminative performance reported in the initial validation studies was substantially attenuated in our older hospitalized cohort and calibration was found to be poor with the model consistently overestimating mortality risk. The results illustrate the importance of testing seemingly accurate prediction models in target populations before applying them in routine practice.

There are plausible reasons for the reduced predictive performance in this external validation study. Firstly, the patients in the present cohort were substantially older (median age was 80 years versus 59 years in the HOMR derivation cohort; see **Table 1**) and less likely to be living independently (66.3% versus 83%).[3] Secondly, unlike the initial validation studies, patients who died during their index hospital admission were excluded. This is likely to be significant because one of the HOMR covariates, the diagnostic risk score, quantifies risk of death based on specific admission diagnoses. High scores associated with diagnoses such as intracerebral haemorrhage and sepsis reflect high risk of death during hospitalization. This risk may diminish significantly when patients survive the initial days of their acute hospital episode. Thirdly, it is unclear whether the diagnostic risk scores, which were derived from a large population of adult patients of all ages, are weighted appropriately for older hospitalized patients. An admission diagnosis of syncope, for example, is assigned a diagnostic risk score of -9 which perhaps reflects its usually benign prognosis in younger adults. Syncope, in older adults however, is associated with reduced survival.[14] Finally, differences in access and organization of primary care between North

America and Ireland may have had an important impact on covariates relating to recent acute hospital care utilization (i.e. ambulance transfers, emergency department visits, readmissions). [15,16]

Our findings are not surprising: the accuracy of predictive models is often substantially lower in new patients compared to the accuracy found in patients of the development population.[17, 18] Rather than simply reject the model, updating methods were used to improve performance in our older patient cohort. In this study, Recalibration in the Large (the simplest updating method where just one parameter of the original model [i.e. the intercept] is adjusted) substantially improved performance. While model revision resulted in further improvements, this more extensive updating method is less ideal because parameter estimates are redeveloped on the data of the validation set (a much smaller sample) and prior information from the larger derivation sample is disregarded.[19]

The performance of the recalibrated HOMR model compares favourably to other validated prognostic models for older hospitalized patients (**Supplementary Appendix S3**). [18, 20-29] However, it is important to emphasize that an updated HOMR model, just like a newly developed model, would require testing of its generalizability, as well as its impact on clinician behaviour and patient outcomes, before it could be recommended for use in routine clinical practice.[30] Even then, because of inherent unwieldiness, it would need to be integrated into hospital information systems to ensure usability for practicing physicians.

The present study has some limitations. Firstly, the HOMR model was applied and updated in a single medical centre where patients were cared for by specialist geriatricians. As discussed, this limits the generalizability of our findings and further validation in other centres is now required. Secondly, we used the model differently to how it was originally designed by excluding patients who died during their index admission. However, we contend that the primary purpose of an accurate 1-year mortality prediction in a hospitalized patient is to help guide decision-making and care-planning *after* the index acute episode when the patient's condition has stabilized.

In conclusion, the exceptional performance of the HOMR model, reported in the North American validation studies, was substantially attenuated in a cohort of older hospitalized patients in a large teaching hospital in Ireland. Nevertheless, the performance of the HOMR model in our older patient cohort was demonstrably good and compares favourably to other validated non-disease specific mortality prediction tools for older people. Updating methods improved performance of the HOMR model but further refinement, validation, as well as clinical impact studies, will be required before the model could be applied confidently in routine practice.

Acknowledgements

Conflict of interest

None

Author contributions

Curtin, O'Mahony, Gallagher: study concept and design. Doyle: data aggregation. Curtin, O'Donnell: determination of covariate values. Dahly, van Smeden: statistic analysis. Curtin, O'Mahony, Gallagher: preparation of manuscript. All authors: critical revision and final approval of manuscript.

Sponsor's role

Curtin and O'Mahony are supported by the European Union's Horizon 2020 research and innovation programme (grant number 634238). The sponsor had no role in the design, conduct or reporting of this study.

References:

1. Steinman MA, Hanlon JT. Managing medications in clinically complex elders: "There's got to be a happy medium". JAMA 2010; 304:1592.
2. Sleeman KE. End-of-life communication: let's talk about death. *J R Coll Physicians Edinb* 2013; 43:197–9.
3. van Walraven C. The Hospital-patient One-year Mortality Risk score accurately predicted long-term death risk in hospitalized patients. J Clin Epidemiol. 2014 Sep;67(9):1025-34.
4. van Walraven C, McAlister FA, Bakal JA, Hawken S, Donzé J. External validation of the Hospital-patient One-year Mortality Risk (HOMR) model for predicting death within 1 year after hospital admission. CMAJ. 2015 Jul 14;187(10):725-33.
5. Irish Coding Standards, Healthcare Pricing Office, January 2016 ii

6. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 2016;35(2):214-26.
7. Clark D, Armstrong M, Allen A et al. Imminence of death among hospital inpatients: Prevalent cohort study. Palliative Medicine 2014;28(6):274-279.
8. Pilotto A, Rengo F, Marchionni N et al. Comparing the prognostic accuracy for all-cause mortality of frailty instruments: a multicentre 1-year follow-up in hospitalized older patients. PLoS One. 2012;7(1):e29090. doi: 10.1371/journal.pone.0029090.
9. Steyerberg EW, Vickers AJ, Cook NR et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21(1):128-38.
10. Harrell, FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Cham: Springer; 2015.
11. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: Chapman & Hall; 1993.
12. Vergouwe Y, Nieboer D, Oostenbrink R et al. A closed testing procedure to select an appropriate method for updating prediction models. Stat Med. 2017;36(28):4529-4539.
13. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2008. ISBN 3-900051-07-0, URL http://www.R-project.org.
14. Soteriades ES, Evans JC, Larson MG et al. Incidence and prognosis of syncope. N Engl J Med 2002;347:878-85.

15. StarfieldB, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q. 2005;83(3):457–502.
16. White B, Carney PA, Flynn J, Marino M, Fields S. Reducing hospital readmissions through primary care practice transformation. J Fam Pract. 2014 Feb;63(2):67-73.
17. Bleeker SE, Moll HA, Steyerberg EW et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 2003;56:826e32.
18. Monacelli F, Tafuro M, Molfetta L et al. Evaluation of prognostic indices in elderly hospitalized patients. Geriatr Gerontol Int. 2017;17(6):1015-1021.
19. Janssen K, Moons K, Kalkman C, et al. Updating methods improved the performance of a clinical prediction model in new patients. J Clin Epidemiol 2008; 61: 76–86.
20. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. JAMA. 2012;307(2):182-92.
21. Teno JM, Harrell FE Jr, Knaus W et al. Prediction of survival for older hospitalized patients: the HELP survival model. Hospitalized Elderly Longitudinal Project. *J Am Geriatr Soc* 2000;48(Suppl 5): S16-24.
22. Walter LC, Brand RJ, Counsell SR, et al. Development and validation of a prognostic index for 1-year mortality in older adults after hospitalization. *JAMA* 2001;285:2987-94.
23. Inouye SK, Bogardus ST Jr, Vitagliano G et al. Burden of illness score for elderly persons: risk adjustment incorporating the cumulative impact of diseases, physiologic abnormalities, and functional impairments. Med Care 2003; 41: 70–83.
24. Fischer SM, Gozansky WS, Sauaia A, et al. A practical tool to identify patients who may benefit from a palliative approach: the CARING criteria. *J Pain Symptom Manage* 2006;31:285-92.

25. Levine SK, Sachs GA, Jin L, et al. A prognostic model for 1-year mortality in older adults after hospital discharge. *Am J Med* 2007;120:455-60.
26. Pilotto A, Ferrucci L, Franceschi M, et al. Development and validation of a multidimensional prognostic index for one-year mortality from comprehensive geriatric assessment in hospitalized older patients. *Rejuvenation Res* 2008;11:151-61.
27. Pilotto A, Rengo F, Marchionni N, et al. Comparing the prognostic accuracy for all-cause mortality of frailty instruments: a multicentre 1-year follow-up in hospitalized older patients. *PLoS ONE* 2012;7:e29090.
28. Dramé M, Novella JL, Lang PO, et al. Derivation and validation of a mortality-risk index from a cohort of frail elderly patients hospitalised in medical wards via emergencies: the SAFES study. *Eur J Epidemiol* 2008;23:783-91.
29. Di Bari M, Balzi D, Roberts AT, et al. Prognostic stratification of older persons based on simple administrative data: development and validation of the “Silver Code,” to be used in emergency department triage. *J Gerontol A Biol Sci Med Sci* 2010;65:159-64.
30. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515-24.

SUPPORTING INFORMATION

Additonal Supporting Information may be found in the online version of this article:

Figure S1. Covariates used to calculate a patient's Hospital-patient One-year Mortality Risk (HOMR) score

Table S2. Re-estimated HOMR model with regression coefficients.

Table S3. Summary of prognostic models used to predict mortality in hospitalized older patients.

Table 1. Baseline characteristics of study participants (and how they compare to original derivation cohort)

Variable	Mean SD	Median [IQR]	(Min, Max)	HOMR derivation cohort
Sex				
Female	800 (56.8%)			61.8%
Male	609 (43.2%)			38.2%
Age	79.3 ± 7.4	80 (74, 85)	(65, 101)	59 (IQR 37 -75)
Living Status*				
Independent	933 (66.2%)			83%
Rehabilitation Unit	33 (2.3%)			0.2%
Homecare	295 (20.9%)			12.1%
Nursing Home	148 (10.5%)			4.5%
Urgency of admission				
Elective	78 (5.5%)			47.4%
ED without Ambulance	498 (35.3%)			25.7%
ED with Ambulance	833 (59.1%)			26.9%
Number of ambulance transfers**	0.3 ± 0.7	0 (0, 0)	(0, 5)	N/A
Admitting Service***				
General Medicine (including geriatric medicine)	1365 (96.9%)			31.4%
General Surgery	3 (0.2%)			11%
Cardiology	17 (1.2%)			6.4%
Orthopedics	8 (0.6%)			8.4%
Gastroenterology/Nephrology/ Neurology	16 (1.1%)			4.9%
ICU admission (directly from emergency department)	3 (0.2%)			7.4%
Home O_2*	0			2.3%
ED Visits**				
0	828 (58.8%)			55.1%
≥1	581 (41.2%)			44.9%
Urgent readmission within 30 days	131 (9.3%)			4.5%
DRS	-1.9 ± 4.8	0 (-1, 0)	(-22, 9)	N/A
CCI****				
0	23.3%			57.8%
1-2	34.2%			21.7%
≥3	42.5%			20.5%

Legend: CCI =Charlson Comorbidity Index; DRS = Diagnostic Risk Score; ED = emergency department; HOMR = Hospital-patient One-year Mortality Risk; ICU = intensive care unit; IQR = interquartile range; N/A = not available; SD = standard deviation. *Prior to index hospitalization. ** In 12 months prior to index hospitalization.*** All patients, after hospital admission, were under the care of the specialist geriatric medicine service. **** Not adjusted for patient age.

Figure 1. Calibration plots of the unadjusted and updated Hospital-patient One year Mortality Risk (HOMR) models: (A) Original HOMR model; (B) Recalibrated model (Recalibration in the Large)

Table 2. Performance of the unadjusted and updated Hospital-patient One-year Mortality Risk (HOMR) models.

	HOMR model	Calibration in the Large	Logistic Recalibration	Model Revision
Intercept	0	-0.42	-0.43	-
Slope	1	1	0.99	-
Residual deviance	1139.96	1107.76	1107.73	1046.55
Df	1409	1408	1407	1389
LRT Chisq p-value	-	<0.001	0.85	-
Brier score (rescaled)	0.15 (0.1 to 0.21)*	0.19 (0.13 to 0.25)	0.19 (0.13 to 0.26)	0.23 (0.18 to 0.31)
Emax	0.103 (0.085 to 0.146)	0.111 (0.03 to 0.225)	0.121 (0.03 to 0.236)	0.017 (0.016 to 0.094)
Eavg	0.058 (0.046 to 0.072)	0.016 (0.01 to 0.028)	0.017 (0.009 to 0.029)	0.008 (0.005 to 0.016)
c-statistic	0.78 (0.76 to 0.81)	0.78 (0.75 to 0.81)	0.78 (0.76 to 0.81)	0.82 (0.8 to 0.85)
* Bootstrapped 95% confidence intervals				

Df = degrees of freedom; LRT = likelihood ratio test; Emax = maximum absolute difference in predicted and calibrated probabilities; Eavg = average absolute difference in predicted and calibrated probabilities.