Automated Identification of Patients with Advanced Illness

Rajdeep Brar, MD[1], Jamie S. Hirsch, MD[1,2], Kevin Bock, MD[1,2], Marsha Meytlis, PhD[1]

[1]Department of Information Services, Northwell Health, Lake Success, NY
[2]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY

Introduction

Accurate identification of patients with advanced illness facilitates alignment of medical therapies with a patient's goals of care, reduces unnecessary or unwanted treatments, and potentially limits extended patient hospitalizations and readmissions. Surveys show that while most patients prefer to die in their home, in reality, most die in the hospital.[1] To facilitate advanced illness patient identification, a machine learning model, which leveraged electronic health record data, was created to predict the probability of a patient having advanced illness. This pilot study was designed to evaluate three classification algorithms for best performance in identifying patients with advanced illness.

Methods

We queried a clinical data repository and obtained a dataset from inpatient stays between January 1, 2017, and December 31, 2018. This dataset contained 45,460 visits from 27,575 unique patients with age $\geq$ 75 years old. Eliminating rows with incomplete data reduced the set to 45,457 visits and 27,574 unique patients. We compared three algorithms which identify advanced illness patients: a baseline algorithm currently used by our health system palliative care team (LACE score $\geq$ 10), a recently published hospital frailty risk score (HFRS),[3] and a locally adapted, random forest modification of the HFRS (HFRS-RF). The LACE score is a risk stratification tool developed to predict risk of 30 day readmission and mortality.[2] The hospital HFRS-based models leverage unsupervised learning and consist of two steps: 1) Ward's hierarchical clustering of patients to identify an advanced illness cluster, and 2) HFRS derivation based upon ICD10 diagnosis codes. The two models differed in how frailty risk score was derived. Model 1 used coefficients from a logistic regression algorithm to derive the HFRS based on previously published work, while model 2 used the inverse of a random forest rank to derive the frailty risk score. Both models were trained on 70%, with the remaining 30% used for testing. In addition to ICD10 codes, we used the number of hospital days as features.

Results

Based on the hierarchical clustering results, we identified the advanced illness cluster as that which had over-representation of diagnoses associated with advanced illness. We ran the three models to see how well they discriminate between patients in the frailty cluster vs. not (present in any of the other clusters). The HFRS-RF model showed the best discriminatory ability as measured by AUC 0.78 (vs. 0.76 for HFRS and 0.60 for LACE >= 10).

We then compared the association of the HFRS-RF and LACE >= 10 with outcomes including long length of hospital stay (LOS) > 10 days, hospital readmission within 30 days of discharge, and being seen by the palliative care service (surrogate for advanced illness in our institution, based on local expert opinion). The out of sample HFRS-RF was computed for visits from January 1, 2019, to May 31, 2019, based on chart diagnoses present from 1 year prior to 24 hours post-admission time. The LACE score was computed as of 24 hours post-admission time. Table 1 shows the respective odds ratios as well as odds ratios adjusted for age, gender, and Charlson Comorbidity Index. The LACE was not adjusted to simulate current use of the score to identify patients with advanced illness by our palliative care team.

Outcome	OR	OR 95% CI	OR Adjusted	OR Adjusted 95% CI	p-value
HFRS-RF - Long LOS	1.44	1.29 - 1.61	1.39	1.24 - 1.56	< .01
LACE >= 10 - Long LOS	1.44	1.37 - 1.52			< .01
HFRS-RF - Readmission within 30 days	1.51	1.35 - 1.68	1.49	1.33 - 1.66	< .01
LACE >= 10 - Readmission within 30 days	2.38	2.27 - 2.50			< .01
HFRS-RF - Seen by Palliative Care	2.03	1.80 - 2.28	1.78	1.57 - 2.00	< .01
LACE >= 10 - Seen by Palliative Care	1.16	1.09 - 1.23			< .01

Table 1

Discussion

The outcomes analysis shows very similar association between long LOS and LACE >=10 and the adjusted HFRS-RF. LACE ≥ 10 is mostly strongly associated with the risk of readmission within 30 days, which is unsurprising as this score was developed to predict that outcome. The adjusted HFRS-RF is more strongly associated with being seen by the palliative care service. The HFRS-RF is a promising and relatively simple scoring method that shows a stronger correlation to being seen by palliative care (and therefore having advanced illness) than the currently used LACE score ≥ 10, and may facilitate improved and more efficient case-detection.

References

1. Bekelman JE, Halpern SD, Blankart CR, et al. Comparison of Site of Death, Health Care Utilization, and Hospital Expenditures for Patients Dying With Cancer in 7 Developed Countries. JAMA. 2016;315(3):272–283. doi:10.1001/jama.2015.18603
2. Van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. CMAJ 2010;182:551–7. 10.1503/cmaj.091117
3. Gilbert T, Neuburger J, Kraindler J, Keeble E, Smith P, Ariti C, et al. Development and validation of a frailty risk score focusing on older people in acute care settings using electronic hospital records. Lancet. 2017;391:1775–1782. doi: 10.1016/S0140-6736(18)30668-8.