A Comparison of Risk Adjustment Models for Hospital Length of Stay

Dr. Tolulope Sajobi, Dr. Mingshan Lu, Jason Jiang, Kelsey Lucyk, Dr. Hude Quan

UNIVERSITY OF CALGARY

the methods hub

tools for better health.

Background

- Budgetary constraints in health care
  - Cost containment
  - Improved efficiency
- Increased research into measures of system performance
- Length of stay is a key indicator of inpatient resource use and hospital efficiency
- Large variation exists in different hospitals
  - Policy decisions, such as budgeting and performance evaluation might account for differences

Background

- Increased interest in risk prediction modelling
  - Patient level information to predict and explain variation and make predictions
  - Predict likelihood of longer hospital stay after adjusting for differences in
    - Resources
    - Case mix
    - Environmental factors
    - Patient attributes
- Risk prediction models for hospital LOS
  - Disease specific models
  - Generic risk prediction models

Background

- Generic risk prediction models useful for health organization and health policy researchers for use at institutional or system level
- Large variety of statistical methods for generic prediction models of LOS
  - OLS regression
  - Binary and ordinal logistic regression
  - Generalized linear regression
  - Time-to-event regression
  - Hierarchical regression

UNIVERSITY OF CALGARY

Background

- Systematic review of literature
- Lu M., T.T. Sajobi, K. Lucyk, D. Lorenzetti, and H. Quan. 2015. “A systematic review of risk adjustment models of hospital length of stay.” Medical Care 53: 355-365.
  - Specifically looking for studies containing risk prediction models of inpatient LOS
  - 37 articles included in final review
  - Majority of articles used OLS regression or generalized regression
- Large variation in predictive accuracy
  - Adjusted R^2: 0.01 – 0.63

Objective

- Compare predictive accuracy of several risk prediction models commonly used in currently literature using a large administrative database

Methods

- Models tested
  - OLS Regression
  - Log-transformed OLS regression
  - Poisson Regression
  - Gamma Regression
  - Hierarchical Regression
  - Log-transformed hierarchical regression
  - Robust regression

UNIVERSITY OF CALGARY

Methods

- Data Source/Study Population
  - Discharge Abstract Database (DAD)
    - All hospital admissions in Alberta from 2008 to 2010
    - Demographic information
      - Age at admission
      - Sex
    - Patient information
      - Admission type
      - Comorbidities/diagnosis
      - Resource intensity weight
      - procedures

Methods

- Statistical analysis
  - Define prolonged LOS as 90 days
  - Separate risk prediction models for patients with:
    - Normal LOS (<=90 days)
    - Prolonged LOS (>90days)
  - Model variables
    - Age at admission
    - Sex
    - Admission type
    - Number of comorbidities
    - Resource intensity weight

Results

Table 1: Inpatient Clinical Characteristics in Discharge Abstract Data

Patient Clinical Characteristics	**Normal LOS* (N_1 =1871779)**	**PLOS* (N_2 = 18174)**	**Total (N =1889976)**
Length of Stay (Mean, SD)	6.46 (10.98)	179.51 (157.72)	8.13 (25.37)
Age at Admission (Mean, SD)	42.70 (28.50)	64.79 (23.16)	42.91 (28.53)
Resource Intensity Weight (Mean, SD)	1.25(2.06)	22.46 (20.57)	1.45(3.55)
Sex (% Female)	57.62	49.75	57.54
Admission Category (n, %)			
Newborn	253906 (13.56%)	181 (1.00%)	254093 (13.44%)
Elective	592782 (31.67%)	6308 (34.71%)	599097 (31.70%)
Urgent	1025091 (54.77%)	11685 (64.30%)	1036786 (54.86%)
Number of Morbidities			
1	382320 (20.43%)	316 (1.74%)	382640 (20.25%)
> 1	1489459(79.57)	17858 (98.26%)	1507336 (79.75%)
Discharge Destination (%)			
Dead/Still Born	51554 (2.75%)	1757 (9.67%)	53312 (2.82%)
Transferred to Other Facilities	315612 (16.86%)	12126 (66.72)	327740 (17.34%)
Discharged Home	1504613 (80.38%)	4291 (23.61%)	1508924 (79.84)

**LOS = length of stay, Normal: (LOS ≤ 90 days); PLOS = LOS > 90 days*

UNIVERSITY OF CALGARY

Results

Table 2: Comparison of Predictive Accuracies of Regression-Based Risk Prediction Models

Risk Adjustment Model	Model R^2(%)	RMSE	MAPE
Normal LOS ($\leq$ 90 days)			
OLS Regression	54.25	7.41	3.65
Log-Transformed OLS Regression	48.10	**0.78**	**0.61**
Poisson Regression	23.12	290.81	5.30
Gamma Regression	0.01	207.07	15.36
Hierarchical Regression	**55.72**	10.84	6.60
Log-Transformed Hierarchical Regression	49.65	0.84	0.65
Robust Regression	52.49	8.17	3.48
Prolonged LOS (> 90 days)			
OLS Regression	57.40	103.62	54.53
Log-Transformed OLS Regression	51.77	0.35	0.26
Poisson Regression	1.86	157.14	175.48
Gamma Regression	0.19	158.45	221.19
Hierarchical Regression	**70.38**	85.95	48.18
Log-Transformed Hierarchical Regression	61.21	**0.32**	**0.24**
Robust Regression	55.49	105.81	52.04

NB: LOS = Length of Stay; MAPE = Mean absolute prediction error; RMSE = Root mean square error; Values in bold fonts correspond to models with better predictive accuracy.

UNIVERSITY OF CALGARY

Conclusion/Discussion

- Hierarchical log transformed linear regression had the best predictive performance
  - Smallest RMSE and MAPE
- Hierarchical linear regression had highest R^2
- We recommend hierarchical linear regression models
  - Consistent with study by Moran et al.
  - Also consistent with the adoption of hierarchical linear regression in commercial risk adjustment software such as DxCG, which is used by Medicare in the U.S.

Limitations

- Threshold of PLOS
    - Choice of 90 days affect predictive accuracy
    - Decided on clinically informed threshold rather than a data driven one
    - Easier to identify patients with LOS over 90 days versus patients with LOS over 30 days
- Measure of predictive accuracy
    - Should models be judged using R^2 or RMSE and MAPE?
- Only included regression models
    - Future research needed on alternative models

Acknowledgement

- MSI Foundation
- AIHS
- Dr. Tolulope Sajobi
- Dr. Mingshan Lu
- Kelsey Lucyk
- Dr. Hude Quan

UNIVERSITY OF CALGARY

Thank You!