HPR130

Comparison of Diagnosis-Based Risk Adjustment Models for Episode-Based Cost

Juyoung Kim, PhD[1]; Minsu Ock, PhD[1]; In-Hwan Oh, PhD[2], PhD; Min-Woo Jo, PhD[1]; Yoon Kim, PhD[3]; Moo-Song Lee, PhD[1]; Sang-il Lee, PhD[1]

[1]Department of Preventive Medicine, University of Ulsan; [2]Department of Preventive Medicine, Kyung Hee University, [3]Department of Health Policy and Management, Seoul National University

CONTACT

Juyoung Kim, RN, MSc, PhD

Department of Preventive Medicine, University of Ulsan College of Medicine
juyoungkim0716@gmail.com

ACKNOWLEDGEMENT

This was a part of PhD thesis of J. Kim at the University of Ulsan College of Medicine, Seoul, South Korea.

REFERENCES

1. OECD. Health at a glance 2019: OECD indicators. Paris (FR): OECD Publishing; 2019.
2. Iezzoni LI. Risk adjustment for measuring health care outcomes. AUPHA; 2013.
3. Health Insurance Review & Assessment Service (HIRA). [Patient sample cohort data]. 2022. Available at: https://opendata.hira.or.kr/op/opc/selectPatDataAplInfoView.do. Korean. Accessed January 8, 2022
4. Health Insurance Review & Assessment Service (HIRA). [KDRG Version 4.2]. Wonju, Gangwondo (KR): Health Insurance Review & Assessment Service; 2018. Korean.

INTRODUCTION

The average growth rate in per capita health spending in OECD countries has recently reached the economic growth rate[1]. The economic burden due to surging health spending stood out the importance of efficiency in the quality of care. For efficiency measurement, risk adjustment was introduced to assure comparability between outcomes[2].

However, there has been no study evaluating the performance of diagnosis-based risk adjustment methods for episode-based costs was not found in South Korea. Therefore, this research aimed to compare the performance of diagnosis-based risk adjustment methods for the episode-based cost in efficiency measurement.

METHODS

We used Korean insurance claims data, the Health Insurance Review and Assessment Service – National Patient Sample (HIRA-NPS)[3]. A separate linear regression model was constructed using 2018 HIRA-NPS, depending on the Major Diagnostic Category (MDC) of the Korean Diagnosis-Related Group (KDRG) (Supplemental Table 1)[4]. Individual models consist of demographic characteristics, types of insurance, institutional types, treatment types, diagnosis-based risk adjustment methods, and episode-based costs.

Following diagnostic-based risk adjustment methods were used to adjust the risk of comorbidities: Refined Diagnosis Related Group (RDRG), Charlson Comorbidity Index (CCI), National Health Insurance Service Hierarchical Condition Categories (NHIS-HCC), and Department of Health and Human Service-HCC (HHS-HCC). We compared model performance using R-squared (R^2), Mean Absolute Error (MAE), and Predictive Ratio (PR). External validity was evaluated using HIRA-NPS in 2017.

RESULTS

We observed improved model performance after adjusting risk for comorbidities in all models (Fig 1). For example, the model including RDRG improved mean adjusted R^2 compared to no risk-adjusted model for comorbidities, resulting in superior performance than CCI (ADRG R^2, 34.2%; CCI R^2, 36.1%; RDRG R^2, 38.5%). In addition, the model performances of two HCCs were superior to the others (NHIS-HCC R^2, 40.6%; HHS-HCC R^2, 41.4%).

The variability of model performance depending on MDC groups was also observed. For example, NHIS-HCC or HHS-HCC showed the highest explanatory power in 13 MDCs, including MDC P (Newborns). On the other hand, RDRG showed the highest adjusted R^2 in 6 MDCs, including MDC C (Eye diseases or disorders) and MDC O (Pregnancy, childbirth, and puerperium), some of which were subjected to the KDRG payment system.

The overall MAEs were the lowest in the model with RDRG (USD 1,178) and the highest no risk-adjusted model for comorbidities (USD 1,241) (Fig 2). The PRs showed similar patterns between models in the following subgroups: age, sex, institutional types, types of insurance, and upper and lower ten percentile of actual costs (Table 1). External validity using the data of 2017 also showed a similar pattern in the model performance (Supplemental Fig 1, 2).

Fig 1. Adjusted R^2 (%) of models

MDC	ADRG	RDRG	ADRG +CCI	ADRG +NHIS-HCC	ADRG +HHS-HCC
B	29.2	35.9	31.7	36.8	35.6
C	35.8	43.1	37.7	47.6	37.8
D	43.6	48.5	44.6	47.4	46.6
E	35.5	38.2	37.0	40.9	40.7
F	48.5	51.0	49.6	51.9	52.0
G	42.5	47.3	43.9	46.0	46.4
H	28.1	30.6	29.4	36.4	38.0
I	53.2	55.2	54.2	55.3	55.4
J	32.3	36.3	36.9	40.6	41.5
K	28.7	34.6	31.0	39.0	38.1
L	25.0	31.7	28.0	33.9	30.7
M	26.4	35.2	28.2	34.3	30.6
N	33.4	40.5	36.4	39.0	37.6
O	30.7	34.2	30.6	30.8	32.2
P	74.4	74.4	74.4	77.1	79.0
R	23.5	30.5	25.4	28.9	29.9
ST	22.8	25.3	27.6	34.3	39.8
UV	15.5	13.6	16.1	22.1	45.6
WXY	19.9	25.1	22.2	29.8	29.3

Fig 2. Mean absolute errors of models

MDC	ADRG	RDRG	ADRG +CCI	ADRG +NHIS-HCC	ADRG +HHS-HCC
B	2,171	2,020	2,159	2,097	2,109
C	571	549	571	546	567
D	462	441	465	469	463
E	1,303	1,238	1,286	1,261	1,276
F	2,135	2,053	2,143	2,157	2,171
G	888	833	884	875	873
H	1,973	1,881	1,976	1,958	1,943
I	918	892	911	907	907
J	1,151	1,091	1,127	1,116	1,111
K	1,067	1,004	1,059	1,021	1,030
L	1,518	1,421	1,508	1,459	1,488
M	1,024	952	1,037	1,016	1,035
N	861	787	858	848	850
O	400	364	399	399	395
P	1,450	1,450	1,460	1,431	1,411
R	2,813	2,580	2,839	2,810	2,764
ST	1,463	1,411	1,426	1,406	1,495
UV	2,130	2,078	2,114	2,078	1,950
WXY	1,535	1,484	1,546	1,487	1,482

Table 1. Predictive ratios of the models

Variables		ADRG	RDRG	ADRG +CCI	ADRG +NHIS-HCC	ADRG +HHS-HCC
Sex	Male	1.000	1.000	1.000	1.000	1.000
	Female	1.000	1.000	1.000	1.000	1.000
Age group	0–2	1.000	1.001	1.000	1.000	1.000
	3–19	1.000	1.126	1.000	1.000	1.000
	20–39	1.000	1.035	1.000	1.000	1.000
	40–59	1.000	1.007	1.000	1.000	1.000
	60+	1.000	0.979	1.000	1.000	1.000
Provider type	Tertiary hospital	1.000	1.000	1.000	1.000	1.000
	General hospital	1.000	1.000	1.000	1.000	1.000
	Hospital	1.000	1.000	1.000	1.000	1.000
Insurance type	Health Insurance	1.000	1.000	1.000	1.000	1.000
	Medical aid	1.000	1.000	1.000	1.000	1.000
Acutal costs	Lower 10 pct	3.700	3.413	3.556	3.413	3.386
	Middle	1.227	1.206	1.219	1.205	1.206
	Upper 10 pct	0.613	0.649	0.627	0.651	0.650

DISCUSSION AND CONCLUSIONS

Firstly, we observed similar performance patterns across the MDCs, as in previous research using DRGs. MDCs related to musculoskeletal (MDC I) or circulatory (MDC F) systems showed higher explanatory powers than the other MDCs, whereas MDC UV (Mental Diseases and Disorders) was the lowest. Secondly, HCCs were preferable methods in efficiency measurement to CCI or RDRG. Average explanatory powers were higher in HCC models than the others. Although RDRG was better in MAE, RDRG does not differentiate complications and comorbidities within the KDRG-based payment system. Thirdly, we could not define episodes and lookback periods (to identify comorbidities) considering seasonal variations of the epidemiological data due to the limitation of the data source. The HIRA-NPS does not have a continuity of follow-up between years. In addition, we still need to conclude the superiority between NHIS-HCC and HHS-HCC. Therefore, further studies should examine the model validity of HCCs with data sources securing continuity. Lastly, our episode costs did not include out-of-pocket costs (i.e., non-payment items) due to the limitation of the Korean payment system. A more precise model could be constructed if the payment system obtains information on out-of-pocket costs.