



Original Research

Prediction of outpatient visits and expenditure under the Universal Coverage Scheme in Bangkok using subscriber's attributes: A random forest analysis

K. Mongkonchoo^{a,*}, H. Yamana^b, S. Aso^c, M. Machida^d, Y. Takasaki^d, T. Jo^b, H. Yasunaga^e, V. Chongsuvivatwong^f, T. Liabsuetrakul^f

^a National Health Security Office, Government Complex, Bangkok, Thailand

^b Department of Health Services Research, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

^c Department of Biostatistics & Bioinformatics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

^d The Partnership Project for Global Health and Universal Health Coverage, Bangkok, Thailand

^e Department of Clinical Epidemiology and Health Economics, School of Public Health, The University of Tokyo, Tokyo, Japan

^f Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand



ARTICLE INFO

Keywords:

Capitation

Health insurance

Outpatient payment

ABSTRACT

Objectives: There is limited evidence on methods to allocate budgets to healthcare providers under capitation schemes. The objective of this study was to construct and test models that predict outpatient visits and expenditure for each healthcare facility using subscriber data from the preceding year.

Study design: We used the database of the Universal Coverage Scheme in Bangkok, Thailand that stores subscriber information and healthcare service utilization data. One-percent and ten-percent random samples of subscribers were selected as training and testing groups, respectively.

Methods: Using data of the training group, we constructed a model using a random forest algorithm to predict outpatient visits and expenditure in 2017 from the 2016 data. The model was applied to the testing group and facility-level predicted number of visits and expenditure were compared with actual data.

Results: The identically-structured training and testing groups consisted of 37,259 and 371,650 subscribers, respectively. Approximately 25% of subscribers utilized outpatient services. The R^2 for models predicting facility-level utilization rate (visits/subscribers) and expenditure per subscriber in 2017 were 0.85 and 0.75, respectively.

Conclusions: The model to predict outpatient visits and expenditure performed well. Such a prediction model may be useful for allocating budgets to healthcare facilities under capitation systems.

1. Introduction

Universal health coverage (UHC) is a global agenda for all countries to provide health services and financial risk protection to all their citizens [1]. The concept of UHC, as stated in the world health report 2010, focuses on three dimensions, namely the health services that are needed, the number of people that need health services, and the costs to whoever must pay for the services [2]. A country's health financing systems must be strengthened to ensure that financial barriers to access health care be removed and that no financial hardship occurs after utilization of health care. Improving universal coverage requires health finance systems that address sufficient funds, pool these funds to spread financial risks, and spend the funds wisely [3].

Financial risk protection for health care has been established in Thailand since 1975. Currently, there are three major health insurance schemes. The Universal Coverage Scheme (UCS), managed by the National Health Security Office (NHSO), covers the majority of the population [4]. Under UCS, patients are provided with basic outpatient services free of charge. Payments are managed by the capitation allocated to health-care facilities. In Bangkok, the capital city of Thailand, the NHSO Region 13 Bangkok (BKKNHSO) receives and manages a budget for outpatient services from the NHSO Headquarter Office through the differential capitation by age structure of subscribers in the province. The budget is then allocated to health-care facilities within the administrative level of the BKKNHSO. However, there have been

* Corresponding author.

E-mail address: khachon.m@nhso.go.th (K. Mongkonchoo).

<https://doi.org/10.1016/j.puhip.2021.100190>

Received 31 August 2021; Accepted 22 September 2021

Available online 30 September 2021

2666-5352/© 2021 The Authors. Published by Elsevier Ltd on behalf of The Royal Society for Public Health. This is an open access article under the CC

BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

informal feedback from health-care facilities that allocation of capitation for outpatient services and their health expenditures is not balanced.

Although methods to pool and spend funds for health are important for health finance systems globally, there is limited evidence on the optimal methods for allocating a budget to individual health-care providers under capitation schemes. Basic demographics such as age and sex have been shown to explain a small proportion of the variation in health-care expenditure [5–8]. Previous studies from high-income countries such as England, Canada, and Italy have shown that more equitable allocation may be possible by building prediction models with more variables including morbidity measures, services provided, health-related quality of life, and socioeconomic status [8–13]. Such methods may be particularly important in resource-limited regions. However, to our knowledge, no such study from middle-income countries have been reported.

Following the development of a population-based database of health-care service data by the BKKNHHSO, a detailed analysis of UCS subscriber data in Bangkok became possible. Therefore, we conducted a study to construct a model that uses subscribers' attributes from one particular fiscal year to predict the number of outpatient visits and expenditure in the subsequent fiscal year. We evaluated the correlations between the predicted and actual number of visits and expenditure for health-care facilities in Bangkok.

2. Methods

This study was approved by an institutional review board and conducted under an agreement with BKKNHHSO to utilize the de-identified data.

2.1. Data source

We used the database of BKKNHHSO which stores data of all health-care services (outpatient services, inpatient services, health promotion and prevention services) provided to UCS subscribers in Bangkok. The service transaction data are submitted to BKKNHHSO from health-care facilities (hospitals, clinics, and health centers). Data on diagnosis, procedures, drugs, examinations, and medical charges are included. In addition to health-care service information, subscriber data, including demographic data and registration of non-communicable chronic diseases, are recorded.

The subscriber data and outpatient service transaction data for the 2016 and 2017 fiscal years (FYs), i.e. from October 2015 to September 2017, were used for this study.

2.2. Participants

Our interest was on predicting service utilization of the full population using models derived from a smaller group of subscribers. To explore the feasibility of the method, we randomly sampled 1% and 10% of subscribers observed through FYs 2016–2017 to obtain training and testing groups, respectively. Subscribers who withdrew from UCS in FY2017 were excluded. The datasets were constructed from the database as subscriber-based aggregated data representing subscriber's conditions, frequency of visits, and annual health expenditure on outpatient services.

2.3. Variables

Service utilization details of a subscriber in FY2016 were used as predictors and the summary of services of the same subscriber in FY2017 were used as outcomes. Subscribers' characteristics included sex, age, home registration (Bangkok resident or migrant), main healthcare facility that they were registered to (main contracting unit), and whether a subscriber ever utilized a service during FY2016 (user or nonuser). Age

was classified by 10-year categories, and those over 80 years were categorized into single group. The chronic disease registration data were used to identify subscriber's underlying non-communicable diseases including cancer, diabetes, hypertension, stroke, asthma, chronic obstructive pulmonary diseases (COPD) and cardiovascular diseases (CVD).

Nineteen diagnoses categorized by International Classification of Diseases, 10th revision (ICD-10) chapters were identified from diagnosis records in the service transaction data in FY2016. The diagnoses and ICD-10 codes are presented in [Supplementary Table 1](#). Diagnosis covariates were categorized into 3 groups: presence of diagnosis records (diagnosed user), history of service utilization during FY2016 but no record of diagnosis (non-diagnosed user), and no record of diagnosis because no services were utilized (non-user).

Medical charges were used to identify utilization of the following 8 services during FY2016: medicines; blood transfusion; examinations; radiology; special diagnostics such as echocardiography; medical equipment (for example, infusion pump and oxygen saturation monitor), procedure and anesthesia; hospital fee; and doctor fee. These medical charges were used as binary variables (resources consumed or not for each type of service). Health promotion and prevention services were summarized into the total number of visits. Inpatient data were used to identify the number of admissions, length of hospitalization, and relative weight based on the Diagnosis-Related Group system.

The two outcome variables of this study were the number of outpatient visits and expenditure in FY2017, which also included the referral outpatient services that were provided at referral hospitals and billed to the main contracting unit of each subscriber.

2.4. Statistical analysis

The data were analyzed using R version 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria, 2017). The 'randomForest' package (version 4.6–14) in R was used to conduct the random forest analyses. Prediction models were constructed at the subscriber level using data from the training group with subscribers' attributes in FY2016 as predictors. The following 43 candidate covariates were included: patient type (user or non-user), sex, age, home registration, non-communicable diseases from the chronic disease register (coded as 7 dummy variables), diagnoses from transaction data (coded as 19 dummy variables), use of services (coded as 8 dummy variables), number of outpatient visits, number of health promotion and prevention services, number of inpatient admission, total length-of-stay in hospital, and total relative weight of inpatient services. The outcome variables were the incremental number of outpatient visits and expenditure in FY2017 compared with those of FY2016. A different model was constructed for each outcome. The number of trees was set at 500 and the number of random candidate variables to be entered was selected by running the *tuneRF* function. We applied the models to the testing group to calculate the predicted number of visits and total expenditure in FY2017. The R script for our prediction is presented in the Supplementary Material.

The subscriber-level data were aggregated into facility-level data to calculate the predicted and actual number of visits and expenditure for each contracting unit which subscribers were registered to. The following facility-level variables were also calculated: number of visits per subscriber (utilization rate), and average expenditure per subscriber. Descriptive statistics of predicted and actual values of these variables were compared with stratification by the type of health-care facility (hospital, clinic, and health center). Predicted utilization rate and average expenditure per subscriber were compared with actual values using correlation analysis.

3. Results

There were 37,259 and 371,650 subscribers selected for the training

and testing groups, respectively. Subscriber characteristics are presented in [Table 1](#). There were no apparent differences in subscriber's attributes between the training and testing groups. The proportions of males and females were similar. The most common age group was 0–10 years. One-fourth of all subscribers sought outpatient services in a year. Among subscribers who sought services, the most common diagnoses were diseases of the respiratory system, diseases of the circulatory system, endocrine, nutritional and metabolic diseases, and diseases of the digestive system. Hypertension was the most prevalent non-communicable disease recorded in the chronic disease register (8%). The majority of subscribers who visited an outpatient department received medicines (84%) and almost half received laboratory tests. Subscriber characteristics aggregated at facility level are presented in [Supplementary Table 2](#).

A summary of service utilization in FY2016 and FY2017 is presented in [Table 2](#). Within the 91,238 users in the testing group, the average frequency of visits was 6.02 per year and the average annual expenditure was 4758.88 baht per user in FY2016. The overall utilization rate in Bangkok was thus 1.48 visits/subscriber/year and the average annual expenditure per subscriber was 1319.60 baht. In FY2017 the average frequency of visits was 6.01 per year and the average annual expenditure was 5371.84 baht per user. The overall utilization rate and annual expenditure per subscriber in FY2017 were 1.48 and 1318.76 baht, respectively. Service utilization aggregated at facility level are presented in [Supplementary Table 3](#).

The random forest algorithm identified the following five variables as the most important predictors for the incremental number of visits

based on an increase in the mean square error: number of visits, hypertension, laboratory tests, medicines, and radiology services. The five most important predictors for incremental expenditure were medicines, laboratory tests, number of health promotion and prevention services, number of outpatient visits, and being diagnosed with hypertension.

The relationships between predicted and actual total number of visits and expenditures in FY2017 are presented in [Fig. 1](#). The actual and predicted utilization rate and average expenditures per subscriber in FY2017 are also presented in the figure. As seen from the graphs, the model using subscribers' attributes in FY2016 predicted the outcomes in FY2017 well. The predicted and actual utilization rates were highly correlated. The overall R^2 values for utilization rate and average expenditure were 0.85 and 0.75, respectively. Within each health-care facility, the R^2 values for utilization rate and average expenditure were 0.67 and 0.50, 0.88 and 0.88, and 0.79 and 0.72 for clinics, health centers, and hospitals, respectively.

The total number of outpatient visits, utilization rate, total expenditure, and average expenditure by different types of health facilities are presented in [Table 3](#). The majority of health-care facilities were clinics, and number of registered subscribers was largest in hospitals. The utilization rate and average expenditure were lower in health centers compared with clinics and hospitals. The means of the predicted and actual values for the number of outpatient visits and total expenditure were similar in clinics whereas the means of predicted values were greater than the actual values in hospitals and health centers.

Table 1
Characteristics of subscribers in the training and testing groups in 2016.

Characteristic	Training group (N = 37,259)		Testing group (N = 371,650)	
	n	(%)	n	(%)
Male	19,432	(52.2)	194,945	(52.5)
Age				
≤10	5,842	(15.7)	58,301	(15.7)
11–20	5,659	(15.2)	57,666	(15.5)
21–30	4,225	(11.3)	41,341	(11.1)
31–40	4,457	(12.0)	43,861	(11.8)
41–50	5,513	(14.8)	54,665	(14.7)
51–60	5,242	(14.1)	51,998	(14.0)
61–70	3,681	(9.9)	37,179	(10.0)
71–80	1,826	(4.9)	17,940	(4.8)
>80	814	(2.2)	8,699	(2.3)
Bangkok resident	28,931	(77.6)	288,423	(77.6)
User	9,582	(25.7)	91,238	(24.5)
Diagnosis for seeking service*				
Diseases of the respiratory system	2,741	(28.6)	27,178	(29.8)
Diseases of the circulatory system	2,202	(23.0)	20,370	(22.3)
Endocrine, nutritional and metabolic diseases	1,985	(20.7)	18,217	(20.0)
Diseases of the digestive system	1,771	(18.5)	16,909	(18.5)
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	1,353	(14.1)	12,637	(13.9)
Underlying non-communicable disease				
Hypertension	3,141	(8.4)	28,816	(7.8)
Cardiovascular diseases	339	(0.9)	3,214	(0.9)
Stroke	328	(0.9)	2,420	(0.7)
Diabetes	1,567	(4.2)	14,472	(3.9)
Cancer	446	(1.2)	3,966	(1.1)
Asthma	365	(1.0)	3,324	(0.9)
Chronic obstructive pulmonary disease	116	(0.3)	1,098	(0.3)
Services utilized				
Hospital fee	8,395	(87.6)	79,490	(87.1)
Medicine	8,102	(84.6)	76,827	(84.2)
Examination	4,299	(44.9)	39,958	(43.8)
Doctor fee	4,044	(42.2)	39,187	(43.0)
Procedure and anesthesia	1,675	(17.5)	14,946	(16.4)
Radiology	1,632	(17.0)	14,533	(15.9)
Special diagnostics	903	(9.4)	7,549	(8.3)
Medical equipment	768	(8.0)	7,163	(7.9)
Blood transfusion	102	(1.1)	796	(0.9)

* The five most frequent categories of disease are listed.

Table 2
Service utilization of subscribers in the training and testing groups in 2016 and 2017.

Variable	Training group (N = 37,259)		Testing group (N = 371,650)	
	mean	(SD)	mean	(SD)
Service utilization in fiscal year 2016				
Number of outpatient visits	1.48	(4.5)	1.48	(4.6)
Annual outpatient expenditure, baht	1,286	(5,495)	1,320	(7,300)
Number of health prevention and promotion services	0.86	(2.4)	0.87	(2.4)
Number of inpatient admissions	0.06	(0.4)	0.06	(0.4)
Total length of hospitalization	0.40	(4.6)	0.38	(4.5)
Total relative DRG weight	0.10	(1.0)	0.10	(1.0)
Service Utilization in fiscal year 2017				
Number of outpatient visits	1.50	(4.7)	1.48	(4.8)
Annual outpatient expenditure	1,344	(5,611)	1,319	(6,112)

DRG, Diagnosis Related Grouping; SD, standard deviation.

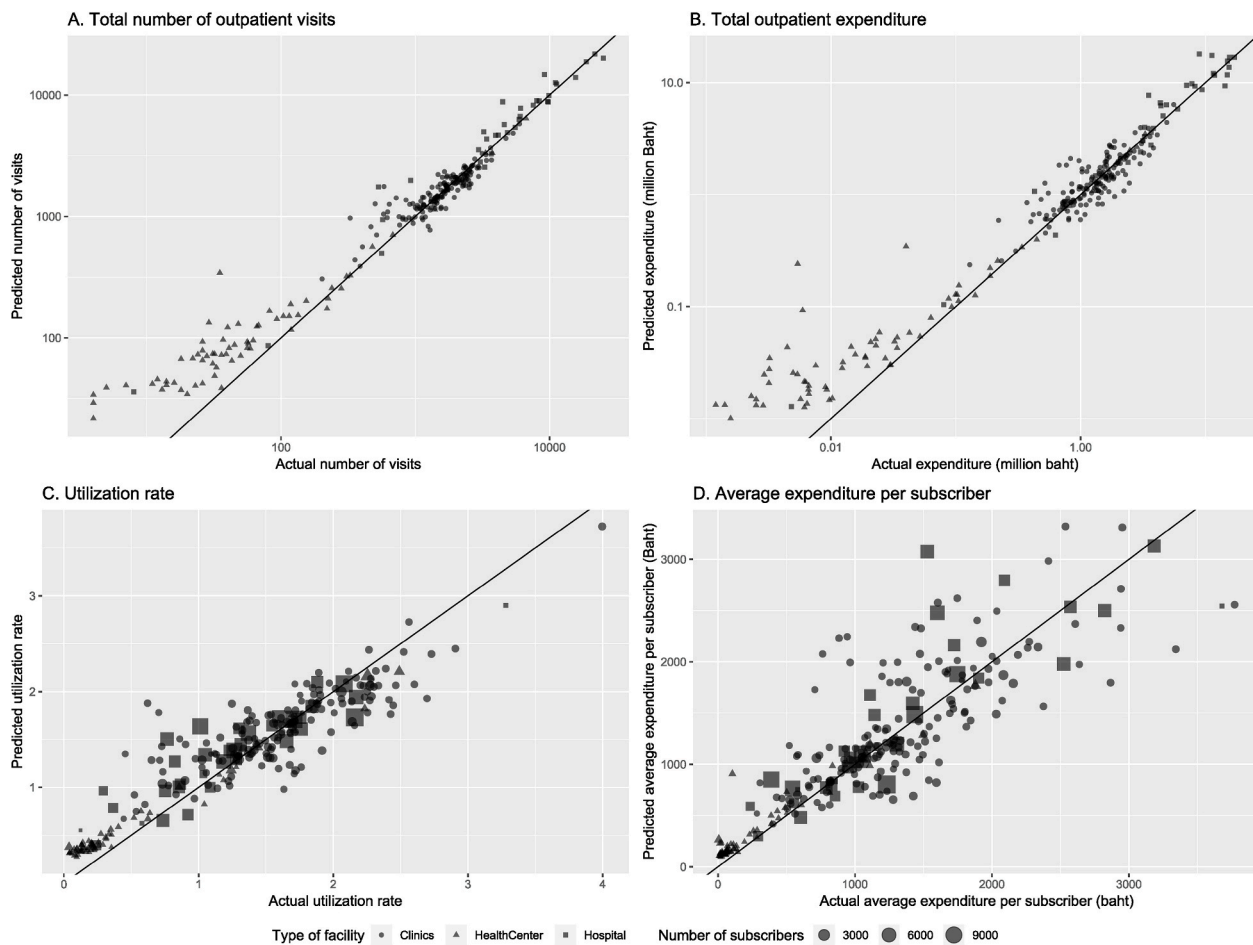


Fig. 1. Relationships between predicted and actual number of visits, expenditure, utilization rate, and average expenditure in the 2017 fiscal year. Utilization rate refers to the number of outpatient visits per subscriber.

Table 3
Predicted and actual facility-level outcome variables in different types of facility.

Variable	Hospitals (N = 35)		Health centers (N = 68)		Clinics (N = 159)		All (N = 262)	
	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual
Number of subscribers		4,834 (2,944)		344 (482)		1,070 (375)		1,384 (1,792)
Total visits	7,183 (5,653)	6,719 (6,040)	361 (960)	344 (1,022)	1,743 (750)	1,754 (869)	2,111 (3,011)	2,051 (3,035)
Utilization rate	1.37 (0.52)	1.25 (0.65)	0.61 (0.47)	0.48 (0.59)	1.63 (0.41)	1.63 (0.55)	1.33 (0.62)	1.28 (0.75)
Total expenditure, million baht	7.0 (5.5)	6.5 (5.2)	0.2 (0.6)	0.2 (0.6)	1.49 (0.85)	1.5 (0.8)	1.91 (2.99)	1.82 (2.79)
Average expenditure per subscriber, baht	1,394 (825)	1,301 (846)	363 (366)	280 (409)	1,396 (593)	1,372 (612)	1,128 (736)	1,079 (767)

Data presented as mean (standard deviation).

4. Discussion

The present study was conducted using a database of all UCS subscribers in Bangkok. Using a 1% sample of subscribers, we built a model with a random forest algorithm that predicted next year's number of outpatient visits and expenditure for outpatient services from information of the previous fiscal year. The model was then applied to a 10% sample of the population. The results were aggregated based on subscribers' registered facilities to predict utilization rate and average expenditure for each facility in FY2017.

The target population of this study was the 3.8 million UCS subscribers in Bangkok. This represents approximately 48% of all people living in Bangkok. Of the randomly selected subscribers in this study, 31% were aged ≤ 20 years and 17% were aged > 60 years. The proportions of children and older adults were higher than those of the entire Bangkok population [14]. This reflects the fact that many of the working-age people in Bangkok belong to other health insurance schemes such as the Social Security Scheme. The data on underlying non-communicable diseases showed that 8% had hypertension and 4% had diabetes. These numbers are consistent with a previous report [15].

Among the study population, 25% of subscribers sought outpatient services while 75% did not. The average frequency of outpatient visits among users was 6 per year and the overall utilization rate was 1.48 visits per subscriber per year. This utilization rate is half the rate among all UCS subscribers in Thailand [15]. There have been discussions on the relatively low utilization rate among Bangkok residents. The reasons to support our findings include a lack of recognition of their main contracting unit and seeking alternative services such as over-the-counter medications at pharmacies [16]. When we stratified the subscribers by the type of health-care facility, a higher utilization rate was seen for clinics compared with health centers and hospitals. In addition to the difference in subscriber background, access to care may be a possible reason for this difference.

In this study, we used a random forest algorithm to predict the number of visits and expenditure from information in the preceding fiscal year. Machine learning methods have been shown to have better predictive ability than regression models, especially when the number of predictors is large and specification of the regression model is difficult [17]. In a preliminary analysis of this study, we also found that the random forest algorithm had better predictive ability compared with a simple regression model (data not shown). The overall prediction results at the facility level were accurate with R^2 values of 0.85 and 0.75 for utilization rate and annual expenditure, respectively. However, the accuracy varied across different types of facilities. The model performed better for hospitals and health centers while the R^2 was lower for clinics (0.67 for utilization rate and 0.50 for expenditure). Services provided at public hospitals and health centers may be more uniform and predictable whereas clinics are more heterogeneous and therefore relatively difficult to predict.

There have been previous studies that have used models to predict service utilization from different countries. In addition to basic demographic information, researchers have included different variables such as underlying disease, region, and socioeconomic status and refined the models [8–13]. Similarly, we identified different diagnoses recorded in FY2016 to describe the underlying conditions. In addition to the transaction data, we utilized the chronic disease register to obtain underlying conditions. Also, transaction data of BKKNHOS enabled detailed evaluation of different types of services provided to each subscriber. To the best of our knowledge, this is the first study conducted in a setting of a middle-income country that used a similar amount of data, if not more, as studies from high-income countries.

The results of this study imply that prediction of the full population may be possible from smaller a smaller group of subscribers. The BKKNHOS is in the process of validating the model for future use in capitation budget calculation. Further improvements to the accuracy of the prediction model may be possible by adjusting the determinants and

comparing the goodness of fit. The data quality is also an important factor. Currently, the BKKNHOS has a strong data validation process, both automatically in the transaction processing system and human-review process, to improve the quality of outpatient service data. Further improvement in data quality can also help to improve the accuracy of the prediction model. Furthermore, the method presented in this study may be expanded to other regions of Thailand or to other countries that currently use only an age structure for capitation budget calculation.

Several limitations of this study must be acknowledged. First, this study was based on one year of prediction. Prediction results may differ by year. However, it should be noted that there was no specific policy change or event that occurred during the study period. Second, the outpatient transaction data was the main source of data in this study. Additional information such as inpatient data, socioeconomic status, and regional characteristics may also contribute to the prediction. Finally, this study was targeted at UCS subscribers in Bangkok. The results may not necessarily be applicable to other population.

5. Conclusion

Using the database of the BKKNHOS, we constructed a model to predict the next year's frequency of outpatient visits and expenditure on outpatient service from subscribers' attributes of the previous year. The model performed well for three different types of healthcare facilities. The prediction model may be useful for allocating budgets to healthcare facilities under the capitation system.

Ethical approval

This research was approved by the Institutional Review Board for Clinical Research, National Center for Global Health and Medicine and conducted under an agreement with National Health Security Office Region 13 Bangkok to utilize the de-identified data and the Ethics Committee of Faculty of Medicine, Prince of Songkla University (REC No. 62-389-18-1).

Funding

This research was supported by the Partnership Project for Global Health and Universal Health Coverage Project (GLO + UHC), a partnership project under the cooperation between Thailand's Ministry of Public Health, National Health Security Office, and the Japan International Cooperation Agency.

Declaration of competing interest

HY1 and TJ have academic affiliations with the Department of Health Services Research, Graduate School of Medicine, The University of Tokyo, which is supported by Tsumura & Company. Tsumura & Company played no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to publish the results.

Acknowledgements

We are grateful for the support of Dr. Jadej Thammatacharee and Dr. Weraphan Leethanakul from National Health Security Office. We thank Mr. Suheim Buya from the Faculty of Science, Prince of Songkla University, Thailand for technical support, Dr. Sachiko Ono from the Department of Eat-loss Medicine, Graduate School of Medicine, The University of Tokyo, Japan for providing assistance with data analysis, and staffs of National Center for Global Health and Medicine, Japan for providing administrative support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.puhip.2021.100190>.

References

- [1] World Health Organization, The World Health Report 2013: Research for Universal Health Coverage. <https://www.who.int/whr/2013/report/en/>. (Accessed 13 May 2020). Accessed.
- [2] World Health Organization, The World Health Report 2010: Health Systems Financing: the Path to Universal Coverage. <https://www.who.int/whr/2010/en/>. (Accessed 13 May 2020). Accessed.
- [3] D.B. Evans, C. Etienne, Health systems financing and the path to universal coverage, *Bull. World Health Organ.* 88 (2010) 402.
- [4] V. Tangcharoensathien, W. Witthayapipopsakul, W. Panichkriangkrai, W. Patcharanarumol, A. Mills, Health systems development in Thailand: a solid platform for successful implementation of universal health coverage, *Lancet* 391 (2018) 1205–1223.
- [5] N. Rice, P.C. Smith, Capitation and risk adjustment in health care financing: an international progress report, *Milbank Q.* 79 (2001) 81–113.
- [6] R.J. Reid, L. MacWilliam, L. Verhulst, N. Roos, M. Atkinson, Performance of the ACG case-mix system in two Canadian provinces, *Med. Care* 39 (2001) 86–99.
- [7] E. Schokkaert, C. Van de Voorde, Risk selection and the specification of the conventional risk adjustment formula, *J. Health Econ.* 23 (2004) 1237–1259.
- [8] J. Dixon, P. Smith, H. Gravelle, S. Martin, M. Bardsley, N. Rice, et al., A person based formula for allocating commissioning funds to general practices in England: development of a statistical model, *BMJ* 343 (2011) d6608.
- [9] L.M. Sibley, R.H. Glazier, Evaluation of the equity of age-sex adjusted primary care capitation payments in Ontario, Canada, *Health Pol.* 104 (2012) 186–192.
- [10] E. Penno, R. Gauld, R. Audas, How are population-based funding formulae for healthcare composed? A comparative analysis of seven models, *BMC Health Serv. Res.* 13 (2013) 470.
- [11] S.L. Brilleman, H. Gravelle, S. Hollinghurst, S. Purdy, C. Salisbury, F. Windmeijer, Keep it simple? Predicting primary health care costs with clinical morbidity measures, *J. Health Econ.* 35 (2014) 109–122.
- [12] M.C. Corti, F. Avossa, E. Schievano, P. Gallina, E. Ferroni, N. Alba, et al., A case-mix classification system for explaining healthcare costs using administrative data in Italy, *Eur. J. Intern. Med.* 54 (2018) 13–16.
- [13] C. Quercioli, F. Nisticò, G. Troiano, M. Maccari, G. Messina, M. Marducci, et al., Developing a new predictor of health expenditure: preliminary results from a primary healthcare setting, *Publ. Health* 163 (2018) 121–127.
- [14] National Health Security Office Region 13 Bangkok, NHSO Annual Report of Fiscal Year, 2017. Available at, <https://bkk.nhso.go.th/AnnualReport60/>.
- [15] National Health Security Office, Annual Report for the Fiscal Year 2017, Available at, <http://stream.nhso.go.th/wp-content/uploads/ebooks/annual-2017/>.
- [16] Boonsuya C, Aimyong N. Public Health Status Survey in Bangkok 2014. Faculty of Public Health, Mahidol University.
- [17] M. Mazumdar, J.J. Lin, W. Zhang, L. Li, M. Liu, K. Dharmarajan, et al., Comparison of statistical and machine learning models for healthcare cost data: a simulation study motivated by Oncology Care Model (OCM) data, *BMC Health Serv. Res.* 20 (2020) 396.