Machine Learning Based Prediction of Non-communicable Diseases to Improving Intervention Program in Bangladesh

Min Hu[1]*, Yasunobu Nohara[2], Yoshifumi Wakata[2], Ashir Ahmed[3,4], Naoki Nakashima[2] and Masafumi Nakamura[5]

[1] Graduate School of Medical Science, Kyushu University, Fukuoka, Japan

[2] Medical Information Center, Kyushu University Hospital, Fukuoka, Japan

[3] Department of Advanced Information Technology, Kyushu University, Fukuoka, Japan

[4] Grameen Communications, Dhaka, Bangladesh

[5] Faculty of Medical Sciences, Kyushu University, Fukuoka, Japan

Abstract

Background: The prevalence of noncommunicable diseases (NCDs) is increasing throughout the world, including in developing countries. An NCD prevention program using information communication technology was implemented for 2 years in Bangladesh. Health checkup data were collected from 16,741 study subjects. However, the effectiveness of the utilized prevention strategy has not yet been evaluated, and some subjects with a risk of NCD have gone undetected.

Objective: This study aimed to improve intervention strategies by analyzing collected data and proposing a cost-effective personalized predictive model to identify subjects predicted to be at future risk of NCD.

Methods: We selected 2,110 subjects who participated in both years of the program and used a machine learning algorithm, gradient boosting decision tree, to build models that would identify subjects who were at risk of future high blood pressure, blood sugar or body mass index (BMI). We used area under the curve (AUC) of receiver operating characteristic curves and cumulative accuracy profile (CAP) curves to evaluate the performance of our models.

Results: Models showed fairly good performance: the BMI model (AUC=0.910) yielded the greatest AUC whereas the BS model (AUC=0.730) yielded the lowest. CAP curves indicated that the BMI model could correctly identify 98.0% of at-risk subjects at only 50% of the total time cost.

Conclusions: Our models represent powerful tools with which to improve the effect of health intervention programs and the effectiveness at which they are performed with limited medical resources.

Keywords

Health informatics; Preventive medicine; Health checkup; Data mining; Developing countries

1 Introduction

The prevention of non-communicable diseases (NCDs) has interested researchers in recent decades. According to the World Health Organization, NCD causes more deaths annually than all other causes combine, making it the leading cause of death globally [1]. A study [2] implementing an NCD prevention program using information communication technology was conducted by Kyushu University and Grameen Communications. A Portable Health Clinic (PHC) package with medical sensors (e.g., blood pressure [BP] monitors, blood glucose meters, weighing scales) was developed to provide health checkups, health risk assessments, and teleconsultations with a remote physician over Skype. Data were collected from 16,741 subjects at 10 locations (including urban and rural regions) in Bangladesh between July 2012 and February 2014 [2].

Health risk assessments can be made via health checkups with the PHC. A health assessment algorithm, called Bangladesh logic (B-Logic), was created based on international diagnostic standards [3, 4, 5, 6, 7] and introduced into all disease management activities of PHC. A set of interventions including consultation with on-site medical staff, consultation with remote physician, and referral to a hospital was designed to detect subjects at risk of NCDs and to provide appropriate intervention.

However, the effectiveness of B-logic as a prevention strategy for NCDs has not yet been evaluated. Although it can identify subjects currently with certain NCDs (e,g, hypertension), the detection of clinical values (e.g., BP) at levels above a given risk threshold, does not negate the possibility that subjects at risk of NCDs in the near future are not accurately detected by B-logic.

An emerging technique in the field of preventive medicine employs data-driven predictive analytics to accurately derive insights from medical records to predict disease progression [8]. For example, Karimi et al. [9] used a decision tree and support vector machine to predict metabolic syndrome in a sample of 2,107 participants. Moreover, Colubri et al. [10] developed a machine learning pipeline for Ebola prognosis prediction that packages the best models into a mobile application available in clinical care settings. Li et al. [11] used a lasso logistic regression model and predicted the risk for colorectal cancer with personal characteristics and fecal immunological test. As popular and effective approaches to predictive analytics, data science and machine learning are highly regarded due to their success in diagnosis, prediction, and choice of treatment [12, 13, 14].

Accordingly, the current study focused on improving our triage algorithm by analyzing collected medical records and by proposing a cost-effective personalized predictive model based on machine learning techniques to detect individual at risk of progressing to NCDs. Specifically, diabetes and hypertension, which are typical NCDs, along with obesity, which is known to be highly related to many NCDs, were selected as the targets for analysis in this study.

2   Methods

2.1 Source of Data/Participants

The PHC as shown in Figure 1 comprises a set of sensor devices, a data transmission system linked to a mobile network, and a data management application.

Health care service with PHC was provided in five rural villages and five factories/offices in Bangladesh. During the visit of the PHC program, we set up PHC devices in each study area, and residents who lived nearby were asked to voluntarily undergo our health checkup. After completing a questionnaire, which asked about liter- acy, occupation, time since the last meal, present symptoms, past diseases, medication, smoking, weight change, exercise, walking speed, eating behavior, sleeping, and the desire to have a healthy lifestyle, subjects underwent a health

Figure 1: The Portable Health Clinic System package.

checkup using the sensor devices in the PHC. Measurements including blood glucose, blood pressure, weight, height, etc. were performed by qualified health care professionals, whereas the other tests were performed by trained staff. Categorized results for four risk groups, which were graded as green (healthy), yellow (caution), orange (affected), and red (emergent) according to the rule based triage algorithm (B-logic) implemented for PHC, were printed and explained to subjects by local staff. For subjects categorized as being at health risk of NCDs (yellow, orange, or red grade), an intervention consisting of a booklet and consultation with on-site staff was provided. For orange and red grade subjects, we also provided a tele-medical intervention (through Skype) with a physician at our call center (in Dhaka, Bangladesh), which is far away from our study areas. The physician could send a tele-prescription via the network, and subjects could then visit their local pharmacy to purchase the prescribed medication. Subjects identified as red-grade were furthermore refereed to the nearest hospital by a remote physician in the Skype tele-consultation [2].

The PHC field study was conducted between July 2012 and March 2014 (first year: July 2012-February 2013; second year: June 2013-March 2014). Some visited areas of the first years are revisited in the second year based on the outcome of first year to accomplish a follow-up for the subjects of first year. In total, 16,741 subjects were assessed in the first health checkup and 2,110 participated in both years of the program. This study was approved by the Ethics Review Committee of Kyushu University Hospital.

2.2 Risk Groups

Figure 2 shows the data selection process, which classified subjects in groups (A-E), as described below. To better understand the ability of PHC as a NCD prevention program to prevent hypertension, diabetes and obesity, we identified subjects who participated in both years of the program (n=2110; group D in Figure 2) from all subjects who participated in both years of the program.

To target individuals at risk of hypertension, diabetes, or obesity, we conducted a series of data selection procedure

to select the risk groups in our study. First we identified subjects in whom values of three checkup item: BP, blood sugar(BS) and body mass index(BMI) had not decreased in the second study year (Group E in Figure 2).

In addition to considering the separate BP, BS and BMI values, we also acknowledged the fact that some subjects would be at greater risk for NCDs due to having high values for more than one of these three items (e.g. both increased BP and BS). Therefore we further subdivided Group E subjects, identifying those in whom values for any two of the three items had risen in 2013(fourth column of Group E; Figure 2). Finally, from the subjects in Group E, we identified four sub-groups of subjectswhose BP, BS, BMI, repectively, were higher than the yellow-grade of B-logic. Since hypertension, diabetes and obesity are common NCDs and a person can have more than one of these diseases at the same, we also created the target risk group who has more than two of above diseases in the second year as show in Figure 2.

These subjects were selected as target groups for NCD prediction in this study. The thresholds of B-logic yellow grade are shown as in the footnote of Figure 1. Numbers of subjects in each group and thresholds for BP, BS, and BMI are shown in the footnote of Figure 2. The color category distribution of target risk groups is close to result of overall subjects based on the result of previous PHC study [2].

In total, 2,110 subjects (Group D) participated in both the 2012 and 2013 health checkups and had complete health checkup data (Table 1).

2.3 Statistical Analysis Method

For identifying the ability of PHC program on preventing hypertension, diabetes and obesity, we calculated the mean of systolic BP, diastolic BP, BS, and BMI of each year. We further conducted paired sample t-test and Cohen's d effect size to compare the mean values obtained in the two years of our study.

The analysis in this study is conducted using Python (version 3.5.2) and its statistical computation package StatsModels (version 0.8.0) and xgboost (version 0.6) [15, 16, 17].

2.4 Model Design and Validation Method

We designed a model to predict NCDs in the four target groups listed in Figure 2 using a machine learning approach known as gradient boosting decision tree (GBDT). GBDT, as proposed by Friedman [18], produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. The ability of a model to predict unknown data is assessed by applying five fold cross-validation [19, 20].

Cross-validation divides the data into five subsets and repeats the performance measure five times. Each time, one of the five subsets is used as the test data, and the other subsets as the training data. The model was trained using the training data and was then tested using the test data (i.e., to test predictions against known values).

Hyper parameter tuning is conducted via grid search cross validation. The final set of hyper parameter tuning for xgboost models is as follows: learning rate=0.01, estimators=800, gamma=0.2, reg alpha=1, subsample=0.8, colsample bytree=0.8, scale pos weight=1, min child weight=3.3, max depth=3.

We estimated the accuracy of the model by calculating metrics such as the area under the curve (AUC) of receiver operating characteristic curve, sensitivity, specificity, and F-measure of the prediction models. We repeated cross-validation 10 times and calculated the mean of each metrics.

Table 1: Comparison of the distribution among Groups B, C and D of Figure 2.

| Checkup Item (X)	NCD Risk Category(Y)*	Number of subjects of Y in X (%) within Group B (n=15,710)	Number of subjects of Y in X (%) within Group C (n=10,342)	Number of subjects of Y in X (%) within Group D (n=2,110)
BP	Green*	11,711 (74.55%)	8,325 (80.50%)	1,728 (81.90%)
	Yellow, Orange, Red*	3,999 (25.45%)	2,017(19.50%)	382 (18.10%)
BS	Green*	14,395 (91.63%)	9,551 (92.35%)	1,988 (94.22%)
	Yellow, Orange, Red*	1,315 (8.37%)	791(7.65%)	122 (5.78%)
BMI	Green*	12,467 (79.36%)	8,499 (82.18%)	1,675 (79.38%)
	Yellow, Orange, Red*	3,243 (20.64%)	1,843(17.83%)	435 (20.62%)
Sex	Male	8,477 (53.96%)	6,257 (60.50%)	1,370 (64.93%)
	Female	7,233 (46.04%)	4,085 (39.50%)	740 (35.07%)
Age	10-19	32 (0.20%)	28 (0.27%)	5 (0.24%)
	20-29	6,827 (43.46%)	5,931 (57.35%)	1,026 (48.63%)
	30-39	3,936 (25.05%)	2,534 (24.5%)	638 (30.24%)
	40-49	2,354 (14.98%)	986 (9.53%)	256 (12.13%)
	50-59	1,334 (8.49%)	451 (4.36%)	106 (5.02%)
	60-69	792 (5.04%)	259 (2.50%)	44 (2.09%)
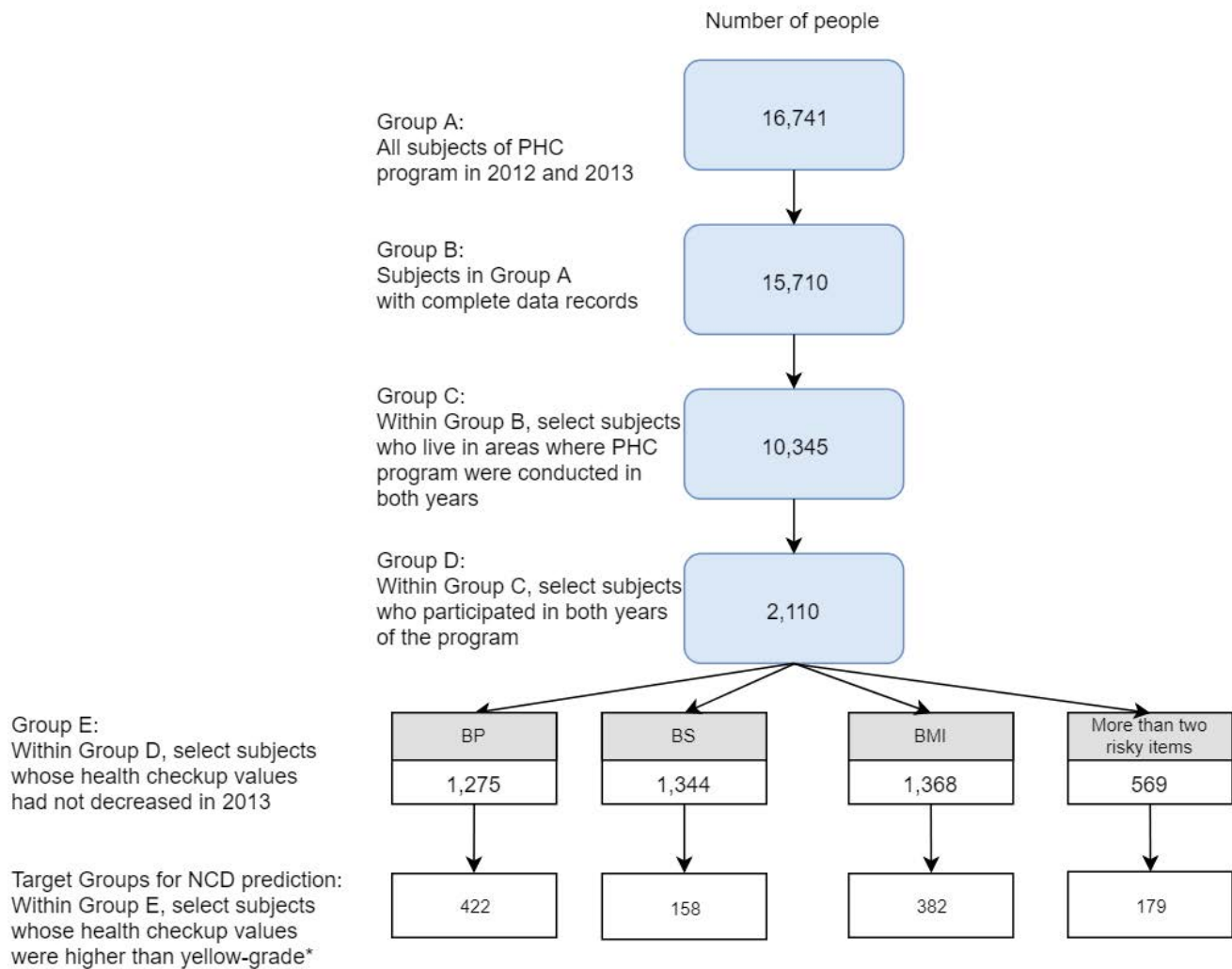
Figure 2: Data selection process used in the current study.

*For BP, the threshold was set as systolic BP ≥ 130 mmHg or diastolic BP ≥ 85 mmHg. For BS, the threshold was set as postprandial BS ≥ 140 mg/dl. For BMI, the threshold was set as BMI ≥ 25 kg/m2. Abbreviations: BMI, body mass index; BP, blood pressure; BS, bloodsugar; PHC, Portable Health Clinic.

2.5 Feature Design

Features from a subjects record were designed to fully cover factors that possibly related to NCDs. Dependent variable in this study is set as a binary variable indicating whether a subject will be grouped as the target group, which is the last row in Figure 2. We use totally 99 variables as explanatory variables in this study.

Basic information and measurement results (49 variables)

The subject profile contains gender, age, site ID (15 sites), site type (rural, suburban, or urban), and checkup. Most of these features are quantitative variables. Measurement items are waist, hip, waist-hip ratio, height, weight, BMI, systolic blood pressure, diastolic blood pressure, blood sugar (BS), BS type (postprandial or fasting), urine protein, urine sugar, urine urobilinogen, pulse rate, arrhythmia, body temperature and SpO2. Each measurement result is classified (four colors) using specific diagnostic rules we implemented for PHC. These colors are also used as explanatory variables.

Interview sheet (50 variables)

Before receiving remote diagnosis by a doctor, the subjects were interviewed by local staff. The interview was conducted using an interview sheet. The interview sheet contains 31 questions about occupation, present symptoms, past diseases, medication, smoking, weight change, exercise, walking speed, eating behavior, sleeping habits, the desire to have a healthy lifestyle, drug allergies, and surgical history. All of these features are categorical variables. Since some questions are given in a multiple choice format and encoded to one of K coding, there are 50 variables designed from interview sheet as predictors.

2.6 Cumulative Accuracy Profile

One way to apply our models in practice is that the models can be used to order the subjects from with most risk

to with least risk, to intervene the risker subjects first. This can lead to a more efficient intervention program with better cost performance. To evaluate our models regarding cost performance in such a situation, we introduced cumulative accuracy profile (CAP) curve in our study.

The application of CAP curve is a popular validation technique currently used in financial practice and is often compared with ROC curves [21]. A CAP curve is similar to the ROC curve; however, the ROC curve compares true positive rate (also known as sensitivity) with false positive rate, whereas the CAP curve relates the true positive rate to the rate of all positives. To the best of our knowledge, this is the first study to apply CAP curves for medical purpose on account of its capability of reflecting a models performance at effectively extracting subjects at risk of NCDs compared with random selection.

As Sobehart proposed [22], to plot CAP curve, subjects are first ordered by model score, from riskiest to safest. For a given fraction x% of the total number of subjects, a CAP curve is constructed by calculating the percentage y(x) of subjects who actually had NCDs in the second year. The y-axis of a CAP curve can be interpreted as a models sensitivity, whereas the x-axis represents the fraction x% of the total number of subjects.

Applying CAP curve herein was a comprehensive and effective way to demonstrate how our model could increase cost performance in the case of PHC program. We thus generated CAP curves for the NCDs risk-predicting models.

3   Results

3.1 Participants' health checkup results across the two study years

In order to identify their demographical difference to the whole group, we compared the number of subjects in Group D with that in Group A in terms of sex, age, BS category, BP category and BMI category as shown in Figure 2.The color category distribution of target risk groups is close to the result of overall subjects based on the result of previous PHC study [2].

The BP, BS and BMI results of subjects who joined in both years are presented in Table 2. In this group, mean systolic BP

decreased from 120.31 mmHg in the first year(2012) to 115.53 mmHg in the second year (2013). Similarly, mean of diastolic BP decreased from 77.38 mmHg (2012) to 76.89 mmHg (2013). On the contrary, BS and BMI both increased in the second year, with BS increasing from 102.55 mg/dl (2012) to 108.82 mg/dl (2013), and BMI increasing from 22.26 kg/m^2 (2012) to 22.66 kg/m^2 (2013). The p-values and effect sizes (Cohen's d) of these differences are listed in Table 2; the differences between years in all four checkup values were significant, with systolic BP having the largest effect size.

In the PHC program, the applied intervention depends on the B-logic color ranking. Therefore, we considered the effects of the intervention to vary in groups of different color and generated Table 2 to show the difference in systolic BP, BS, and BMI based on subjects' color rankings in the first year. There was a significant decrease in systolic BP for all color rankings. For BS, there was a significant increase in the green group and a significant decrease in the orange and red groups.BMI increased significantly in the green and yellow groups.

3.2 Prediction Result

Prediction metrics

We generated a GBDT model to predict four target groups. Table 3 shows the model's prediction performance in terms of AUC, sensitivity, specificity, f-measure, and obtained accuracy. The models showed fairly good performance. The BMI model (AUC=0.910) yielded the greatest AUC among the four items whereas the BP model (AUC=0.730) yielded the lowest. To unify with the result of the CAP curves, sensitivity, specificity, f-measure, and accuracy were generated by setting the median of predicting scores from training data as the threshold for each model. The BMI model had the highest AUC among the four models.

CAP curve result

Figure 3 shows the model selection CAP curves compared with random selection. In order to improve the efficiency of the intervention program, we can consider a scenario in which the subjects are intervened not immediately after health checkup, but first ranked from with most risk to with

Table 2: Comparison of mean BP, BS and BMI values from 2 years of health checkups (n=2,110).

	Systolic BP (mmHg)	Diastolic BP (mmHg)	BS (mg/dl)	BMI (kg/m2)
2012 value (mean)	120.31	77.38	102.55	22.26
2013 value (mean)	115.53	76.89	108.82	22.66
Student's T test's p-value to compare two means	p<0.001	p<0.05	p<0.001	p<0.001
Effect Size (Cohen's d)	-0.294	-0.046	0.166	0.12

Abbreviations: BMI: Body Mass Index; BP: Blood Pressure; BS: Blood Sugar.

Table 3: Results of evaluation of models predicting risk of specific health indicators.

Health checkup item	AUC	Sensitivity	Specificity	f-measure	Accuracy
BP	0.73	0.756	0.568	0.44	0.608
BS	0.761	0.823	0.526	0.213	0.548
BMI	0.91	0.98	0.607	0.522	0.675
Two or three risk items combined	0.876	0.972	0.544	0.283	0.581

Abbreviations: BMI-Body mass index; BP- Blood pressure; BS-Blood sugar.

least risk based on the health checkup result of total population. Risker subjects are prioritized on intervention in order to take the most advantage of limited medical resource.

In this scenario, the accuracy of the model will determine the efficiency of the intervention program. By comparing with randomly selecting subjects for intervention with model selecting intervention, we can better understand how well the model performs on improving the efficiency of the intervention with certain medical resource.

In Figure 3, each point on the model CAP curve indicates the percentage of correctly identified subjects at risk of NCDs (Y -axis) with the percentage of time cost (X-axis) via intervening in the subjects; this is presented sequentially from the subject with the highest to the lowest prediction score. Also in Figure 3, as a comparison of the models in terms of CAP curves, we highlight the percentage of correctly identified at-risk subjects, with 50.

The four BP, BS, BMI and combined risk prediction models correctly detected 75.6%, 82.3%, 98.0% and 97.2% of subjects at risk of NCDs respectively, with half of the time cost.

4 Discussion and Conclusion

In recent years, a great deal of preventive healthcare research utilized predictive models with electronic health record data [8, 9, 10, 12, 13, 14], and studies focusing on predictive models for NCDs have also been reported [23, 24].

For example, Wu et al. [23] compared three different machine learning approaches to predict heart failure diagnosis more than 6 months before the actual date of clinical diagnosis. Guo et al. [24] fitted logistic models to predict overweight and obesity in adulthood from BMI in childhood.

In this study, we attempted to improve intervention strategy through analyzing collected data, and proposing a cost-effective personalized predictive model to predict the subjects with risk of NCD in the future. Our model should be a powerful tool to improve the effect of health intervention program as well as effectively performing intervention with limited medical resource.

Ours is the first study to use data mining to build models that predict subjects most likely to have hypertension, diabetes or obesity at the same time using health checkup data from a developing country.

4.1 Interpretation

Our model performed well at predicting subjects who are likely to have obesity, with an AUC at 0.910. An AUC over 0.9 is considered excellent performance [25]. The BMI model to detect obesity performed better than the others models, which is likely due to the fact that the subjects did not consider obesity as great of a health risk as hypertension and diabetes. Specifically, they were not encouraged to change their lifestyle as much as subjects with hypertension or diabetes, thus outcomes 1 year later were more easily predicted by the model.

The BP and BS models had AUCs of 0.730 and 0.761, respectively, and performed less well than the BMI model. However, evaluation using CAP curves on the test data revealed that these models were able to correctly identify 75.6%, 82.3% and 98.0% of subjects with hypertension, diabetes, and obesity, respectively, with half of the total subjects selected. In other words, these values represent the percentage of work that can be done at 50% of the time cost. We used 50% as a benchmark value in our study to compare the performance between different models in terms of cost reduction. However, results associated with any other values could be generated quite simply using CAP curves, as shown in Figure 4.

4.2 General Applicability and Limitation

Our models were fitted and evaluated with the data collected mostly from suburban factories and rural villages of Bangladesh. To assess its general applicability, we must further evaluate these models with data collected from other developing countries and areas.

The validation done in this study is internal due to the limitation of our sample size. Also, the health intervention program was only conducted for two years, which leads to the fact that our model is built only to predict the risk in the next year. In reality, it is necessary to apply the models to a longer window like the risk of 5-10 years later.

4.1 Conclusion

In this study, we designed models that can predict, with high accuracy, the incident rate of NCDs such as hypertension,
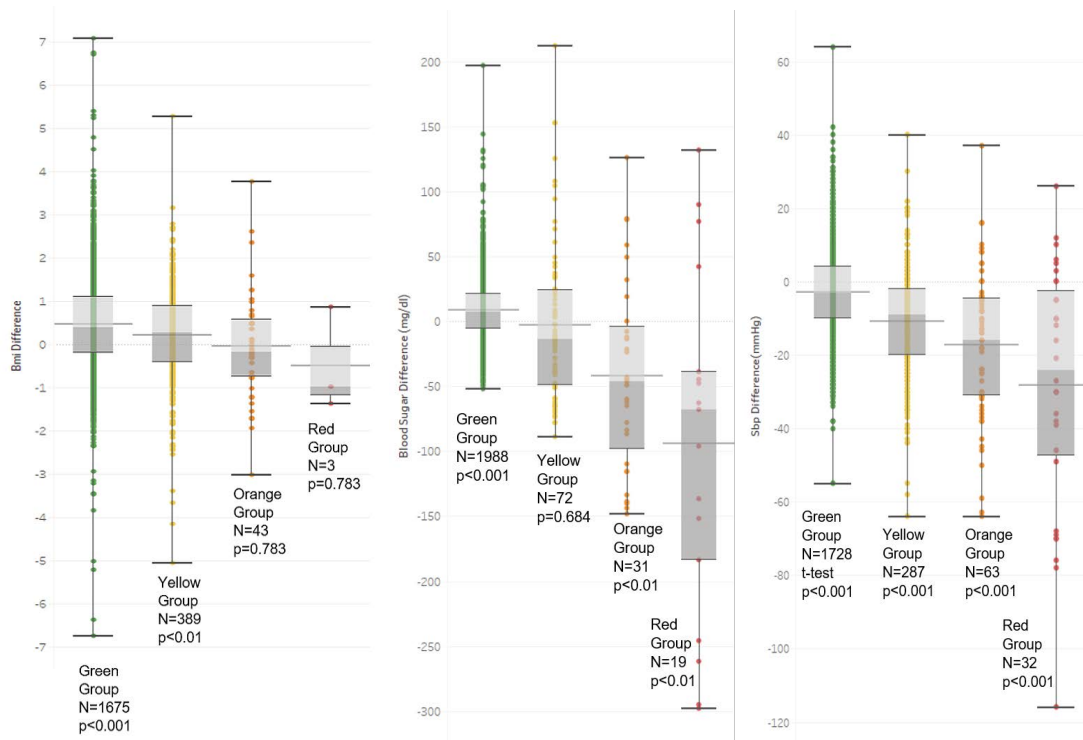
Figure 3: Comparison of 2 years of health checkup data in subject groups based on color ranking. BP, BS and BMI values based on color ranking of a health assessment algorithm (Bangladesh logic or "B-logic").

Abbreviations: BMI-Body mass index; BP- Blood pressure; BS-Blood sugar; SBP-systolic blood pressure.

(a) BP.

(b) BS.

(c) BMI.

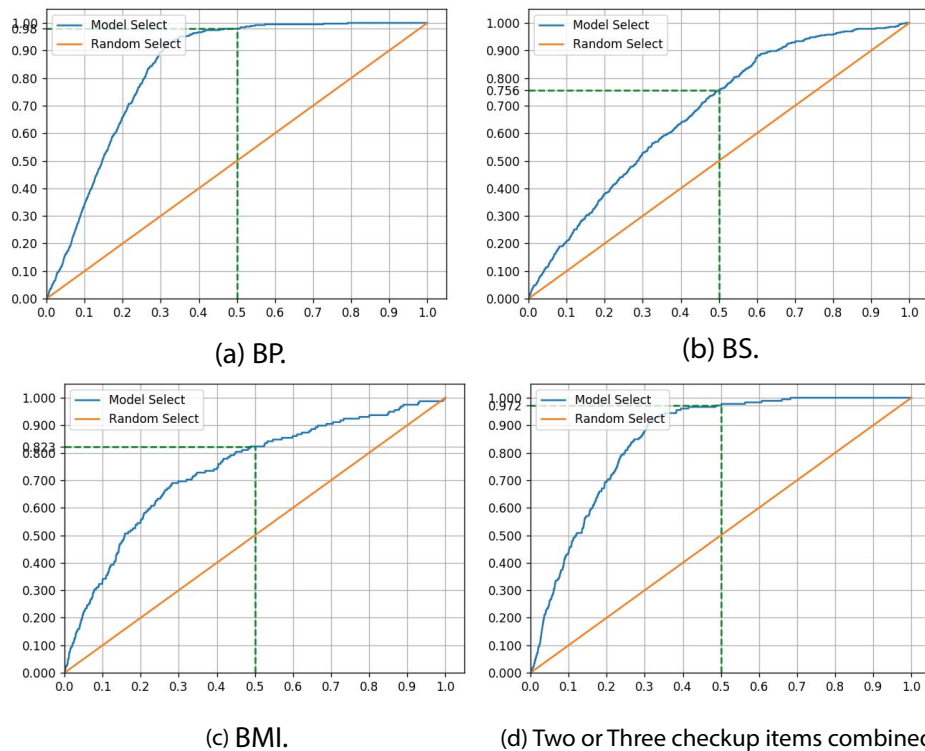(d) Two or Three checkup items combined

Figure 4: Cumulative accuracy profile curves of the four models predicting the risk of hypertension, diabetes, and obesity via assessment of BP, BS, BMI and more than two of these health checkup item combined, in comparison with random selection.

Abbreviations: BMI: Body Mass Index; BP: Blood Pressure; BS: Blood Sugar; SBP: Systolic Blood Pressure.

diabetes or obesity among patients in following year. The models represent powerful tools that could be used to improve the effect of health intervention programs as well as to promote the effective conduction of interventions in situations with limited medical resources. This is achieved via the data-driven detection of patients who are at risk of NCDs.

In the future, we aim to use the results of this study and apply our models to our established health intervention programs in developing countries.

5   Author Contribution

Naoki Nakashima, Yasunobu Nohara, Yoshifumi Wakata, Ashir Ahmed and Masafumi Nakamura contributed to the design and implementation of the research, Min Hu performed the analysis of the results and contributed to the writing of the manuscript.

6   Disclosure Statement

N/A

7   Ethics and Consent

This study is approved by the Ethics Review Committee of Kyushu University Hospital.

8   Funding Information

9   Paper Context

The prevalence of noncommunicable diseases is increasing throughout the world, including in developing countries. A noncommunicable disease prevention program using information communication technology was implemented for two years in Bangladesh and collected data from 16,741 subjects. This study aimed to improve intervention strategies by analyzing collected data and proposes a cost-effective personalized predictive model to identify subjects predicted to be at future risk of noncommunicable disease.

References

[1]   Alwan A. Global status report on non-communicable diseases 2010. Geneva, Switzerland: World Health Organization; 2011.

[2]   Nohara Y, Kai E, Ghosh PP, Islam R, Ahmed A, Kuroda M, et al. Health checkup and tele-medical intervention program for preventive medicine in developing countries: verification study. J Med Internet Res. 2015; 17: e2.

[3]   International Diabetic Federation. Metabolic syndrome criteria. 2006.

[4]   World Health Organization. Waist circumference and waist-hip ratio: Report of a WHO expert consultation. Geneva: World Health Organization; 2008.

[5]   National Heart, Lung and Blood Institute. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. In: The Evidence Report. Bethesda, MD: National Heart, Lung and Blood Institute; Sep 1998: 1-228.

[6]   Cifkova R, Erdine S, Fagard R, Farsang C, Heagerty AM, Kiowski W, et al. Practice guidelines for primary care physicians: 2003 ESH/ESC hypertension guidelines. J Hypertens. 2003; 21: 1779-1786.

[7]   International Diabetes Federation. Global guideline for type 2 diabetes. 2012.

[8]   Sharafoddini A, Dubin JA, Lee J. Patient Similarity in Prediction Models Based on Health Data: A Scoping Review. JMIR Med Inform. 2017; 5: e7.

[9]   Karimi-Alavijeh F, Jalili S, Sadeghi M. Predicting metabolic syndrome using decision tree and support vector machine methods. ARYA Atherosclerosis. 2016; 12: 146-152.

[10] Colubri A, Silver T, Fradet T, Retzepi K, Fry B, Sabeti P. Transforming Clinical Data into Actionable Prognosis Models: Machine-Learning Framework and Field-Deployable App to Predict Outcome of Ebola Patients. PLoS Negl Trop Dis. 2016; 10.3: e0004549.

[11] Li W, Zhao LZ, Ma DW, Wang DZ, Shi L, Wang HL, et al. Predicting the risk for colorectal cancer with personal characteristics and fecal immunochemical test. Medicine. 2018; 97: e0529.

[12] Mi X, Ikeda H, Nakazawa F, Matsuoka H, Kataoka E, Hamaya S, et al. Prescription Prediction towards Computer-Assisted Diagnosis for Kampo Medicine. In: Computer Application Technologies (CCATS) International Conference; 2015 Aug 31; Matsue, Japan. IEEE; 2015. p. 126-131.

[13] Soguero-Ruiz C, Mora-Jimenez I, lvarez JLRA, Hindberg K. Feature selection using Kernel Component Analysis For Early Detection of Anastomosis Leakage. Proceedings of the 2nd International Workshop on Pattern Recognition for Healthcare Analytics. 2014 Aug 24; Stockholm, Sweden; 2014.

[14] Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell. 2015; 163: 1079-1094.

[15] Oliphant TE. Python for Scientific Computing. Comput Sci Eng. 2007; 9: 10-20.

[16] Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. Proceedings Proceedings of the 9th Python in Science Conference. 2010 June 28-June 3; Austin, Texas; 2010.

[17] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 August 13-17; San Francisco, California, USA; 2016.

[18] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001; 29: 1189-1232.

[19] Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. J Med Internet Res. 2016; 18: e323.

[20] Kurasawa H, Hayashi K, Fujino A, Takasugi K, Haga T, Waki K, et al. Machine-Learning-Based Prediction of a Missed Scheduled Clinical Appointment by Patients With Diabetes. J Diabetes Sci Technol. 2016; 10: 730-736.

[21] Irwin RJ, Irwin TC. Appraising Credit Ratings: Does the CAP Fit Better than the ROC. Int J Finance Econ. 2013; 18: 396-408.

[22] Sobehart JR, Keenan SC, Stein R. Benchmarking Quantitative Default Risk Models: A Validation Methodology. Moodys Investors Service. 2000.

[23] Wu J, Roy J, Stewart WF. Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. Med Care. 2010; 48: S106-S113.

[24] Guo SS, Wu W, Chumlea WC, Roche AF. Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. Am J Clin Nutr. 2002; 76: 653-658.

[25] El Khouli RH, Macura KJ, Barker PB, Habba MR, Jacobs MA, Bluemke DA. Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast. J Magn Reson Imaging. 2009; 30: 999-1004.