



Artificial intelligence universal biomarker prediction tool

Yoshiyasu Takefuji¹

Accepted: 25 November 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Through experiencing cardiopulmonary arrest, an artificial intelligence universal biomarker prediction tool was developed to help patients understand improvement in the trends of their disease. PyPI tool handles two biomarkers, hbA1c for diabetes and NP-proBNP for heart failure, to predict the next hospital visit. Predicting improvement in disease is a great hope for patients.

Keywords Cardiopulmonary arrest · Medication and exercise · hbA1c and NT-proBNP

On October 22, 2022, the author experienced a cardiopulmonary arrest that lasted for 25 min. He was diagnosed with asymptomatic ischemia with diabetes mellitus. Three instances of coronary artery stenosis were identified, with an ejection fraction in the mildly abnormal range. Two stents were subsequently inserted using a catheter. He was hospitalized for 20 days after ICU. Two MRI scans with and without contrast showed no necrosis or damage to the heart.

Two physicians, a cardiologist and a diabetologist, provided drug therapy and suggested drug therapy and daily walking for 30 min in a row as aerobic exercise. Medication and the 30-minute walk are part of routine.

At each visit to the hospital, the two physicians show him the results of his blood tests, but they do not explain in detail how much his heart disease and diabetes have improved. They explained that the numbers in the table are black and red, with red being out of the normal range. The table can show improvement, but cannot detail projections for the next hospital. Patients want to improve their disease with prospects and hopes for the future. Detailed predictions motivated the author to create a system for patients.

Two determinants were chosen based on the advice of two physicians: NT-proBNP for cardiac improvement and hbA1c for diabetes improvement. N-terminal pro-brain natriuretic peptide (NT-proBNP) is the gold standard biomarker for HF

diagnosis and therapy monitoring [1]. Glycated hemoglobin (hbA1c) is considered the gold standard for predicting glycaemia-associated risks for the microvascular and macrovascular complications of diabetes mellitus [2].

The data.csv is composed of seven determinants such as “day”, “hbA1c”, “NT-proBNP”, “degree1”, “degree2”, “y1” and “y2”. “day” determinant indicates the date of hospital visit. Two determinants such as hbA1c and NT-proBNP are biomarkers to be predicted. Columns 4 (degree1) and 5 (degree2) show the degree of polynomial regression for the first and second biomarkers, respectively. Columns 6 (y1) and 7 (y2) set the range of vertical axis for the first and second biomarkers, respectively. The range of y1 is from 5 to 9 and that of y2 is from 0 to 450. data.csv must be created by the patient or physician.

The PyPI phope [3] universal biomarker prediction tool allows users to modify the names of two biomarkers with data.csv file. In the proposed linear trend model, the trend line is the unique line that minimizes the sum of squared deviations from the time-series data for individual patients which is one of the best predictors as far as we know. This is because each individual patient is different. In other words, integrating patients is not a meaningful predictor of patient trends.

Błaziak et al. reviewed the AI approach to guiding the management of heart failure patients [4]. They introduced Machine learning (ML) methods to offer valuable tools for distinguishing between heart failure (HF) patients and individuals without HF, with many current models utilizing heart rate variability for HF detection. While AI and ML techniques have been primarily utilized by physicians and

✉ Yoshiyasu Takefuji
takefuji@keio.jp

¹ Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan

professionals, no tools have been introduced for patients. This paper introduces a universal tool for predicting biomarker values for patients' subsequent hospital visits with small datasets.

The tool in question differs from traditional machine learning methods such as deep learning. This is primarily due to the limited dataset, a characteristic of biomarkers such as HbA1c and NT-proBNP. The HbA1c level is largely influenced by the blood glucose levels of the most recent month, as it depends on the interaction between blood glucose concentration and erythrocyte lifespan. Given the continuous turnover of erythrocytes, it's estimated that an HbA1c value represents glucose exposure as follows: 50% from the preceding 30 days, 40% from the previous 31–90 days, and 10% from the previous 91–120 days [5]. As a result, fluctuations in blood glucose levels in the most recent month have a more pronounced impact on the HbA1c level than changes that occurred 2 or 3 months prior [5]. The proposed tool with professional advices is designed for use with small datasets, in contrast to traditional machine learning methods which typically require larger datasets.

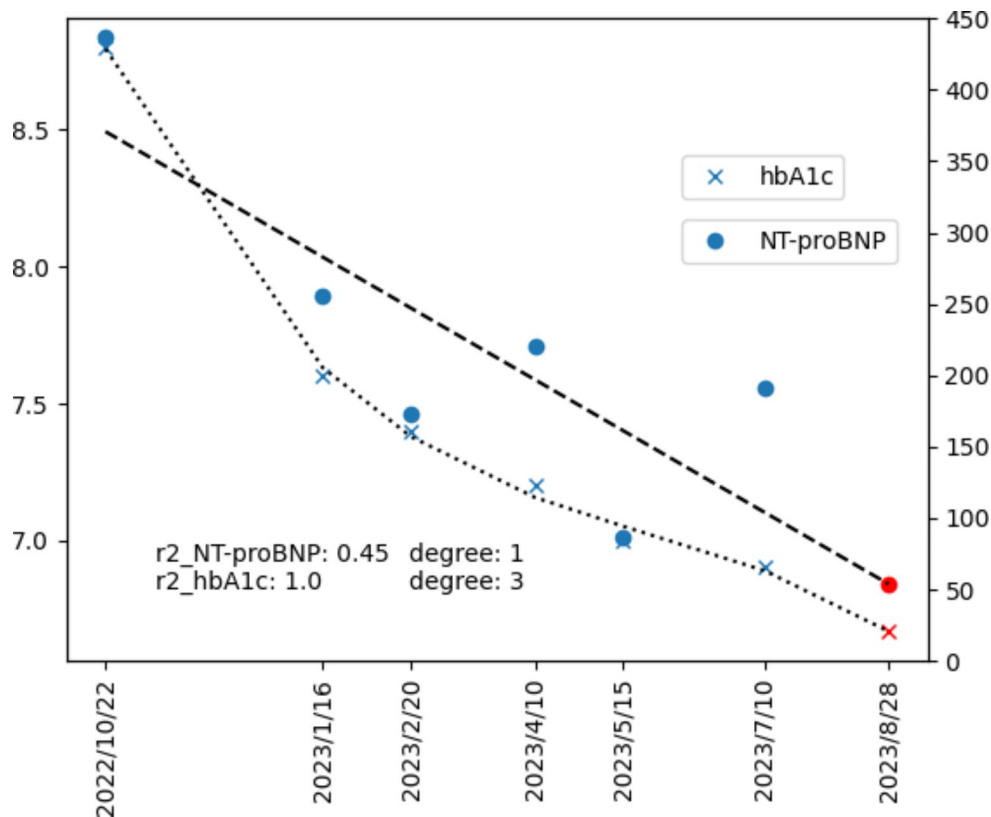
The cardiac neurohormone known as N-terminal pro-B-type natriuretic peptide (NT-proBNP) is secreted when there's an increase in ventricular stress [6]. It serves as an indicator of mortality and cardiovascular incidents in patients suffering from ailments such as type 2 diabetes and heart failure. Nonetheless, the correlation between

NT-proBNP concentrations and the timing of stress is intricate. Although it's accurate that NT-proBNP concentrations can escalate due to sudden cardiac stress, these levels can also be affected by other elements like inflammation, body mass index (BMI), and kidney function. Consequently, while NT-proBNP can offer crucial insights about the heart's condition, it might not accurately reflect the recent weekly stress in comparison to the previous month. It's always advisable to interpret NT-proBNP levels along with other clinical data and under the supervision of a medical professional.

BNP is released in a rhythmic manner as a prohormone when subjected to stress related to volume and pressure. The active hormone of BNP has a half-life of approximately 22 minutes [7]. On the other hand, NT-proBNP has a longer half-life, ranging from 60 to 120 min. For patients with severe heart failure, more frequent testing of NT-proBNP may be necessary. This is because these patients are at a higher risk and their condition may change rapidly. The proposed tool can be used with small datasets for prediction.

To ensure accurate measurements with professional advices, it's feasible to obtain only one data point for these biomarkers approximately every month at most. This study utilizes a limited dataset of these biomarkers for prediction purposes, a task that conventional machine learning techniques are unable to perform due to a small dataset. The tool we're proposing leverages linear regression, a method

Fig. 1 Result of phope with data. csv



frequently employed for trend prediction in scenarios with small datasets.

Figure 1 shows the result of phope with data.csv. The phope tool generated two graphs of NT-proBNP with 'o' and hbA1c with 'x' respectively. It allows users to modify any two biomarker names. Red dots with 'o' and 'x' indicate predicted values from two polynomial regressions for the next hospital visit. Two lines such as bold and thin dotted lines are polynomial regressions. In other words, the left vertical axis shows the hbA1c values and the right vertical axis shows the NT-proBNP values with predicted regression lines.

In Fig. 1, r^2_{hbA1c} indicates R-squared value for hbA1c and $r^2_{\text{NT-proBNP}}$ for that of NT-proBNP. R-Squared value is a statistical measure about the goodness of fit of a regression model. R-squared value allows users to identify the best polynomial regression model. The higher the number, the better the fit. In other words, the higher R-squared value, the better the predicted value.

For the next hospital visit, 53.6 with degree=3 is predicted for NT-proBNP and 6.7 for hbA1c with degree=3 respectively as shown in Fig. 1. Predictive values are very useful to patients and provide strong motivation and hope. Calculated coefficients are generated on the screen. R-squared value indicates the fitting between the predicted model and raw data. Without modifying the program, users are allowed to change the degree of polynomial regressions with given data in data.csv file.

As of October 23, 2023, the phope app has been downloaded 5668 times. This number suggests its popularity worldwide.

Acknowledgements Not applicable.

Author contribution YT completed this research and wrote the program and this article.

Funding Grant and award info: This research has no fund.

Data sharing statement Not applicable.

Declarations

Competing interests The author has no conflict of interest.

Ethical approval Information, institution(s) and number(s): Not applicable.

Patient involvement Not applicable.

References

1. Bellagambi FG, Petersen C, Salvo P et al (2021) Determination and stability of N-terminal pro-brain natriuretic peptide in saliva samples for monitoring Heart Failure. *Sci Rep* 11:13088. <https://doi.org/10.1038/s41598-021-92488-2>
2. Klein KR, Buse JB (2020) The trials and tribulations of determining HbA_{1c} targets for Diabetes Mellitus. *Nat Rev Endocrinol* 16:717–730. <https://doi.org/10.1038/s41574-020-00425-6>
3. PyPI. phope. <https://pypi.org/project/phope/> Access on 18 April 2023
4. Błaziak M, Urban S, Wietrzyk W, Jura M, Iwanek G, Stańczykiewicz B, Kuliczkowski W, Zymlński R, Pondeł M, Berka P, Danel D, Biegus J, Siennicka A (2022) An Artificial Intelligence Approach to guiding the management of Heart Failure patients using predictive models: a systematic review. *Biomedicines* 10(9):2188. <https://doi.org/10.3390/biomedicines10092188>
5. Wang M, Hng TM (2021) HbA_{1c}: More than just a number. *Aust J Gen Pract.* ;50(9):628–632. <https://doi.org/10.31128/AJGP-03-21-5866>. PMID: 34462769
6. Samad M, Malempati S, Restini CBA (2023) Natriuretic Peptides as biomarkers: Narrative Review and considerations in Cardiovascular and respiratory dysfunctions. *Yale J Biol Med* 96(1):137–149. <https://doi.org/10.59249/NCST6937> PMID: 37009194; PMCID: PMC10052602
7. Rodolaki K, Pergialiotis V, Sapantzoglou I, Theodora M, Antsaklis P, Pappa K, Daskalakis G, Papapanagiotou A (2023) N-Terminal Pro-B type natriuretic peptide as a predictive biomarker of Bronchopulmonary Dysplasia or Death due to Bronchopulmonary Dysplasia in Preterm neonates: a systematic review and Meta-analysis. *J Pers Med* 13(9):1287. <https://doi.org/10.3390/jpm13091287> PMID: 37763055; PMCID: PMC10532602

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.