A Clinical Decision Support System for Diabetes Patients with Deep Learning: Experience of a Taiwan Medical Center

Ting-Ying Chien  
Yuan Ze University

Hsien-Wei Ting  
Yuan Ze University

Chih-Fang Chen  
MacKay Memorial Hospital

Cheng-Zen Yang (czyang@saturn.yzu.edu.tw)  
Yuan Ze University

Chong-Yi Chen  
Yuan Ze University

Research Article

Keywords: Bidirectional long short-term memory (Bi-LSTM), clinical decision support system (CDSS), deep learning, diabetes mellitus, electronic health record, oral hypoglycemic agents

Posted Date: August 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-778775/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background: Diabetes mellitus (DM) is a major public health problem worldwide. It involves dysfunction of blood sugar regulation resulting from insulin resistance, inadequate insulin secretion, or excessive glucagon secretion.

Methods: This study collated 971,401 drug usage records of 51,009 DM patients. These data included patient identification code, age, gender, outpatient visit dates, visit codes, medication features (including items, doses, and frequencies of drugs), HbA1c results, and time of testing. We applied a random forest (RF) model for feature selection and implemented a regression model with a bidirectional long short-term memory (Bi-LSTM) deep learning architecture. Finally, we used the root mean square error (RMSE) as the evaluation index for the prediction model.

Results: After data cleaning, the data included 8,729 male and 9,115 female cases. Metformin was suggested to be the most important feature by the RF model, followed by glimepiride, acarbose, pioglitazone, glibenclamide, glipizide, repaglinide, nateglinide, sitagliptin, and vildagliptin. The model performed better when using the past two seasons in the training data than with additional seasons. Further, the Bi-LSTM model performed better than the support vector machine (SVM) model.

Conclusion: This study found that the Bi-LSTM model is a well processing architecture in a clinical decision support system (CDSS), which assists physicians in decision-making, and increasing the number of seasons had a negative impact on the model performance. In addition, this study showed that the most important drug was metformin, which is recommended as the first-line oral hypoglycemic agent (OHA) in DM patients.

Introduction

Diabetes mellitus (DM) is a major public health problem worldwide. It involves dysfunction of blood sugar regulation resulting from insulin resistance, inadequate insulin secretion, or excessive glucagon secretion [1]. There are two types of DM. Type 1 DM is usually due to an autoimmune disorder and involves the destruction of pancreatic beta-cells. Type 2 DM is caused by impairment of glucose regulation due to the malfunction of pancreatic beta cells or insulin resistance [1]. Treatment using oral hypoglycemic agents (OHA) for type 2 DM may have negative side effects, such as hypoglycemia. Therefore, it is crucial to ensure the safety and efficacy of OHA usage [1-7].

Clinical decision support systems (CDSSs), which integrate electronic health records (EHRs) and expert knowledge, have improved the decision-making of physicians and medical caregivers [8-11]. Many methods have been developed for CDSSs, including linear/logistic regression, support vector machines (SVMs), decision trees, random forest (RF), rough sets, and trajectory methods [12-18]. Glycated hemoglobin (HbA1c) is extensively studied in these approaches because it is a good indicator of DM control. DM patients with higher HbA1c measures are more likely to experience renal diseases, macrovascular events, cardiovascular diseases, retinopathies, skin ulceration/gangrene, and high
A well-controlled HbA1c value plays an important role in DM management. CDSSs have been shown to be effective in supporting HbA1c control. For example, O’Connor et al. showed that the HbA1c of DM patients significantly improve when the physicians use a CDSS compared with when physicians do not use a CDSS ($p < 0.01$). Moreover, 94% of physicians using the CDSS were satisfied for this application and physicians continued to use the CDSS for more than one year without research funding support. [19]

Recently, deep learning methods have dramatically improved different fields of medical care and research [20]. They have also been used as the core methods to build the CDSS [21]. For example, convolutional neural networks (CNNs) are used to process image data and recurrent neural networks (RNNs) are used for sequential pattern problems [21, 22]. Sun et al. proposed a method to predict blood sugar levels at four intervals, namely 15, 30, 45, and 60 minutes, using the long short-term memory (LSTM) model and the bidirectional-LSTM (Bi-LSTM) model [23]. Therefore, we devise a CDSS using a Bi-LSTM model with HbA1c as the outcome index for managing OHA usage. The structure of the proposed CDSS, the LSTM model, and the Bi-LSTM model are shown in Figure 1. (Figure 1)

**Materials And Methods**

We collated 971,401 drug usage records of 51,009 diabetes mellitus (DM) patients from January 2012 to December 2014 (12 seasons) and 313,165 laboratory records of 74,792 DM patients in a medical center from January 2012 to June 2015 (14 seasons). These data included patient identification code, age, gender, outpatient visiting dates, visiting code, medication features (included items, doses, and frequencies of drugs), HbA1c results, and testing time. The data were combined and cleansed. Twelve seasons of data and 17,844 DM patients were included in this study. The data were evaluated with five-fold cross-validation (training data = 80% and testing data = 20%) (Figure 2). We applied an RF model with mean square error (MSE) for feature selection where higher means decrease MSE indicated more important parameters [24-26]. OHA dosages and codes were collected. This study was approved by the Institutional Review Board (IRB) of the MacKay Memorial Hospital (IRB approval number: 15MMHIS143e).

We implemented the Bi-LSTM structure using PyTorch running on two personal computers equipped, respectively, with Ubuntu 16.04 and Ubuntu 18.04. The GPU environments were NVIDIA GeForce GTX 980 and GTX 1080 Ti. We used Grid search to adjust the model parameters: epoch (100); batch size (32, 64, 128); hidden layer neuron numbers of Bi-LSTM (32, 64, 128); dropout rate (0, 0.2) between Bi-LSTM; optimizers AMSBound [27] and Adadelda [28]; and learning rate (AMSBound used 0.001 and the Adadelda used 1.0). We used early stopping (patience = 30) and L2 normalization (weight decay = 0.0005) to solve the problem of overfitting and applied gradient clipping (clip norm = 5) to prevent exploding gradient problems. We also implemented a support vector regression (SVR) model using Rgtsvm in R to compare with the Bi-LSTM model [17].

To evaluate the models, we used root mean square error (RMSE)
We applied Pearson’s chi-squared test and the student t-test for data analysis. The statistical analysis was conducted using SPSS version 19.0 (SPSS Inc., Chicago, IL, USA). Statistical significance was defined as $p < 0.05$.

**Results**

Of the included 17,844 cases, 8,729 (49.0%) were male and 9,115 (51.0%) were female. The mean age was 62.3 years old (SD = 11.9) overall, 60.4 (SD = 11.8) for males, and 64.2 (11.7) for females. The 45- to 64-year-old age had the most cases (8,507 cases), followed by those aged above 65 (6,966 cases). The mean Hba1c was 7.6% (SD = 1.7). There were 13,346 cases with Hba1c higher than 6.5% and 4,498 cases whose Hba1c were less than 6.5% (Table 1).

| Demographic data of training cases |
|-----------------------------------|
| Sex                               | Male        | Female      | All         |
| Case number (%)                   | 8,729 (49.0)| 9,115 (51.0)| 17,844      |
| Mean age (SD)                     | 60.4 (11.8) | 64.2 (11.7) | 62.3 (11.9) |
| Age rank                          |             |             |             |
| < 25 years old                    | 22          | 20          | 42          |
| 25–44 years old                   | 732         | 373         | 1,105       |
| 45–64 years old                   | 4,531       | 3,976       | 8,507       |
| ≥ 65 years old                    | 2,882       | 4,084       | 6,966       |
| Unknown age                       | 562         | 662         | 1,224       |
| Mean Hba1c (SD)                   | 7.6% (1.8)  | 7.6% (1.7)  | 7.6% (1.7)  |
| Hba1c ≤ 6.5% (%)                  | 2,337 (26.8%) | 2,161 (23.7%) | 4,498 (25.2%) |
| Hba1c > 6.5% (%)                  | 6,392 (73.2%) | 6,954 (76.3%) | 13,346 (74.8%) |

The data included 11 types of drugs. Compound drugs glimepiride (25,719), pioglitazone (25,720), and vildagliptin (25,726) were combined with metformin dosages of 500, 850, and 1000 mg, respectively.
Nateglinide had dosages of 60 and 120 mg.

The most important feature was metformin with mean decreased MSE = 171.8, followed by glimepiride (156.6), acarbose (151.8), pioglitazone (148.1), glibenclamide (143.7), gliclazide (114.1), repaglinide (93.3), nateglinide (80.6), sitagliptin (74.0), and vildagliptin (21.9) (Table 2).

Table 2

| Oral hypoglycemic agents (OHA) mean decreased mean square error (MSE), dosages, and codes |
|---|---|---|---|
| OHA          | Mean Decrease MSE | Item code(s) | Dosage    |
| Metformin    | 171.8             | 25703        | 500 mg    |
| Glimepiride  | 156.6             | #25709/25719 | 2 mg      |
| Acarbose     | 151.8             | 25721        | 100 mg    |
| Pioglitazone | 148.1             | #25722/25720 | 15 mg     |
| Glibenclamide| 143.7             | 25708        | 5 mg      |
| Gliclazide   | 114.1             | 25713        | 30 mg     |
| Repaglinide  | 93.3              | 25701        | 1 mg      |
| Nateglinide  | 80.6              | $25712/25714 | 60 mg/120 mg |
| Sitagliptin  | 74.0              | 25718        | 100 mg    |
| Vildagliptin | 71.1              | #25724/25726 | 50 mg     |
| Linagliptin  | 21.9              | 25727        | 5 mg      |

# These three compound drugs are all combined with metformin. Glimepiride (25719), pioglitazone (25720), and vildagliptin (25726) have 500, 850, and 1000 mg of metformin added, respectively.

§ Nateglinide has two dosages: 60 and 120 mg.

MSE: mean square error

This study treated every season as ground truth from 2013 Q1 to 2015 Q1 and constructed nine datasets, each having a different sample size. For example, the dataset of 2014 Q4 had 12,677 and 3169 cases as training and test samples, respectively. Using other data as independent factors, we designed three kinds of models. The first used two seasons of data to predict drug usage of the third season. For example, model 9 (2015 Q1) used 2014 Q3 and 2014 Q4 to predict 2015 Q1. The other two types of models used three/four seasons to predict the drugs of the fourth/fifth seasons.
This study also evaluated differences in Hba1c between seasons. For example, we calculated the differences in mean Hba1c between 2015 Q1 and 2014 Q4 (0.87%), 2014 Q3 (0.98%) and 2014 Q2 (1.09%). We found that longer time distances had greater differences in Hba1c (Table 3).

We compared Bi-LSTM and SVM in the two-, three-, and four-season models. The RMSE of both two-season models (Bi-LSTM = 1.05 ± 0.07 and. SVM = 1.05 ± 0.17) was the best, followed by the three-season models (Bi-LSTM = 1.12 ± 0.03 and. SVM = 1.10±0.25) and four-season models (Bi-LSTM = 1.16 ± 0.04 and SVM = 1.09 ± 0.21). The sensitivity and specificity of the two-season Bi-LSTM model was and 0.68 ± 0.05.

The sensitivity of the Bi-LSTM models was not significantly different to each other (two seasons: 0.88 ± 0.03, three seasons: 0.88 ± 0.02, four seasons: 0.89 ± 0.02). The sensitivity of the SVM models gradually decreased non-significantly (two seasons: 0.83±0.16, three seasons: 0.80±0.21, four seasons: 0.77±0.23), but performed worse than the Bi-LSTM models.

The specificity of the Bi-LSTM models gradually decreased as the included seasons increased (two seasons: 0.68 ± 0.05, three seasons: 0.64 ± 0.05, four seasons: 0.59 ± 0.04). The specificity of the SVM models was not significant for any approach (two seasons: 0.69±0.32, three seasons: 0.71 ± 0.31, four seasons: 0.71 ± 0.30). According to the MCC evaluation, there were no significant differences between the six models. The two-season Bi-LSTM model (0.39 ± 0.06) had the shortest run time, followed by the three-season (0.47 ± 0.09) and four-season (0.52 ± 0.06) Bi-LSTM models. The SVM models had significantly longer run times (two seasons: 3.39±0.64, three seasons: 5.18 ± 1.07, four seasons: 5.36 ± 1.17) than the Bi-LSTM models (Table 4).
Discussion

Studies have found that higher Hba1c is linked with increased risk of complications in DM patients [3]. A physician–pharmacist collaboration is useful for OHA adjustment to manage Hba1c owing to physician knowledge and experience [29]. However, it is a challenge to leverage the knowledge of these experts. The current study found that Bi-LSTM models performed better than SVM models for a CDSS to support physicians' decision-making related to OHA adjustment for DM patients. Many CDSS and classification
models have used SVM and other artificial intelligence technologies [17, 30-33]. Although the models had similar RMSE, the Bi-LSTM models were much faster than the SVM models. We were able to successfully incorporate expert knowledge into the CDSS.

We also found that increasing the number of seasons used in the prediction negatively impacted accuracy and RMSE. Physicians reference the most recent Hba1c value to adjust OHA dosage and may choose to maintain the dosage if the Hba1c value is only slightly higher than 7% for the first time. Thus, it is not necessary to reference three or more seasons of Hba1c data, as validated by our experimental results.

We calculated the importance of these drugs and used RF to translate this information into the CDSS [24-26, 30, 32]. The most important drug was metformin, which is recommended as first-line treatment OHA in various situations for DM patients. This drug improves lipids and inflammatory markers and reduces cardiovascular events, but may be contraindicated for patients with mild to moderate chronic kidney disease. Recent research indicates that metformin requires caution in these kinds of DM patients [5]. Sulfonylureas is an important DM drug. We found that the glimepiride (2nd) glibenclamide (5th), and gliclazide (6th) are also important OHA for DM patients [2, 4]. Acarbose is a popular OHA with the same side effects as those of metformin by direct comparison [6]. Pioglitazone is an important DM drug that may reduce Hba1c and improve both metabolic syndrome and nonalcoholic fatty liver disease/nonalcoholic steatohepatitis [7]. It also has a side effect of weight loss for some patients, which is sometimes treated as a benefit. Glimepiride, pioglitazone, and vildagliptin were all combined with metformin (Table 2).

This study has some limitations and areas for extension. Yanase et al. reported that low HbA1c is linked with frailty and suspected malnutrition in elderly type 2 DM patients [34]. Around 25% of cases in this study had Hba1c < 6.5%, indicating that the DM control was too strict for some patients, potentially leading to malnutrition or hypoglycemia. Our CDSS defined “good control” as Hba1c ≤ 7. Although our approach worked well, future versions could be enhanced by considering 6.5 ≤ Hba1c ≤ 7.

**Conclusion**

In recent years, diabetes has become one of the most threatening chronic diseases in many countries. Our study proposed a deep-learning approach to provide physicians with a useful reference by which to improve medication prescription in order to prevent waste in medical resources. We analyzed the medication records and HbA1c test results provided by hospitals. After pre-processing, deep learning neural networks were employed to predict the HbA1c level of diabetic patients, which can assist physicians in prescribing. Due to the time-series nature of the patient medication records and HbA1c test records, we decided to use recurrent neural network (RNN) and Bi-LSTM models as the main models in our network. The RMSE between the model-predicted HbA1c values and the actual HbA1c measurements of patients was overall smaller than that of the SVR model, and hence the Bi-LSTM model can predict the
future HbA1c values of patients more accurately. This allows physicians to determine whether the condition of the patients is gradually improving in the expected direction.

**Abbreviations**

Bi-LSTM: bidirectional long short-term memory  
CDSS: clinical decision support system  
CNNs: convolutional neural networks  
DM: diabetes mellitus  
EHRs: electronic health records  
HbA1c: glycated hemoglobin  
LSTM: long short-term memory  
OHA: oral hypoglycemic agents  
RF: random forest  
RMSE: root mean square error  
RNNs: recurrent neural networks  
SVMs: support vector machines  
SVR: support vector regression

**Declarations**

**Acknowledgements:**

This study was supported by the MacKay Memorial Hospital (MMH-E-10403) and the Ministry of Science and Technology (MOST109-2627-H-155-001), Taiwan. The authors would like to thank the Mackay Memorial Hospital for providing the dataset used in our study.

**Ethics approval and consent to participate**

This study was approved by the Institutional Review Board (IRB) of the MacKay Memorial Hospital (IRB approval number: 15MMHIS143e) and the experimental protocol was established, according to the ethical guidelines of the Helsinki Declaration.
 Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analysed during the current study are not publicly available due to our hospital's policy and the approval of IRB but are available from the corresponding author on reasonable request.

Author Contributions:

T.Y. Chien and C.Z. Yang designed the study; C.F. Chen collected data; C.F. Chen and H.W. Ting provided support regarding knowledge of medicines; T.Y. Chien and C.Z. Yang provided support regarding information technology knowledge; C.Y. Chen constructed the decision support system; T.Y. Chien and H.W. Ting wrote the manuscript; C.Z. Yang made the decision to publish. All authors have read and approved the manuscript.

Conflicts of Interest:

The authors declare no conflict of interest. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Blair M: Diabetes Mellitus Review. Urologic nursing 2016, 36(1):27-36.
2. Chen C-F, Ting H-W, Yang C-Z, Yeh H-I, Hung T-C: Effectiveness of Brand-Name and Generic Versions of Glimepiride for Diabetes Mellitus Care: Experience at a Medical Center in Taiwan. International Journal of Gerontology 2019, 13(2):116-120.
3. Gorst C, Kwok CS, Aslam S, Buchan I, Kontopantelis E, Myint PK, Heatlie G, Loke Y, Rutter MK, Mamas MA: Long-term Glycemic Variability and Risk of Adverse Outcomes: A Systematic Review and Meta-analysis. Diabetes care 2015, 38(12):2354-2369.
4. Kalra S, Bahendeka S, Sahay R, Ghosh S, Md F, Orabi A, Ramaiya K, Al Shammar S, Shrestha D, Shaikh Ket at: Consensus Recommendations on Sulfonylurea and Sulfonylurea Combinations in the Management of Type 2 Diabetes Mellitus - International Task Force. Indian journal of endocrinology and metabolism 2018, 22(1):132-157.
5. Sanchez-Rangel E, Inzucchi SE: Metformin: clinical use in type 2 diabetes. Diabetologia 2017, 60(9):1586-1593.
6. Gu S, Shi J, Tang Z, Sawhney M, Hu H, Shi L, Fonseca V, Dong H: Comparison of glucose lowering effect of metformin and acarbose in type 2 diabetes mellitus: a meta-analysis. *PloS one* 2015, 10(5):e0126704.

7. DeFronzo RA, Inzucchi S, Abdul-Ghani M, Nissen SE: Pioglitazone: The forgotten, cost-effective cardioprotective drug for type 2 diabetes. *Diabetes & vascular disease research* 2019, 16(2):133-143.

8. Légat L, Van Laere S, Nyssen M, Steurbaut S, Dupont AG, Comu P: Clinical Decision Support Systems for Drug Allergy Checking: Systematic Review. *Journal of medical Internet research* 2018, 20(9):e258.

9. Cho I, Bates DW: Behavioral Economics Interventions in Clinical Decision Support Systems. *Yearbook of medical informatics* 2018, 27(1):114-121.

10. Kilsdonk E, Peute LW, Jaspers MW: Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. *International journal of medical informatics* 2017, 98:56-64.

11. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, Samsa G, Hasselblad V, Williams JW, Musty MD et al: Effect of clinical decision-support systems: a systematic review. *Annals of internal medicine* 2012, 157(1):29-43.

12. Breiman L: *Random Forests*. *Machine Learning* 2001, 45(1):5-32.

13. Ting HW, Chien TY, Lai KR, Pan RH, Wu KH, Chen JM, Chan CL: Differences in Spontaneous Intracerebral Hemorrhage Cases between Urban and Rural Regions of Taiwan: Big Data Analytics of Government Open Data. *International journal of environmental research and public health* 2017, 14(12).

14. Tseng TL, Huang CC, Fraser K, Ting HW: Rough set based rule induction in decision making using credible classification and preference from medical application perspective. *Computer methods and programs in biomedicine* 2016, 127:273-289.

15. Ting HW, Wu JT, Chan CL, Lin SL, Chen MS: Decision Model for Acute Appendicitis Treatment With Decision Tree Technology-A Modification of the Alvarado Scoring System. *Journal of the Chinese Medical Association : JCMA* 2010, 73(8):401-406.

16. Chien TY, Lee ML, Wu WL, Ting HW: Exploration of Medical Trajectories of Stroke Patients Based on Group-Based Trajectory Modeling. *International journal of environmental research and public health* 2019, 16(18).

17. Chan1 C-L, Chen C-L, Ting H-W, Phan D-V: An Agile Mortality Prediction Model: Hybrid Logarithm Least-Squares Support Vector Regression with Cautious Random Particle Swarm Optimization. *International Journal of Computational Intelligence Systems* 2018, 11(1):873 - 881.

18. Teixeira AL, Leal JP, Falcao AO: Random forests for feature selection in QSPR Models - an application for predicting standard enthalpy of formation of hydrocarbons. *J Cheminform* 2013, 5(1):9. doi: 10.1186/1758-2946-1185-1189.

19. O’Connor PJ, Sperl-Hillen JM, Rush WA, Johnson PE, Amundson GH, Asche SE, Ekstrom HL, Gilmer TP: Impact of electronic health record clinical decision support on diabetes care: a randomized trial. *Annals of family medicine* 2011, 9(1):12-21.
20. LeCun Y, Bengio Y, Hinton G: Deep learning. Nature2015, 521(7553):436-444.
21. Ting HW, Chung SL, Chen CF, Chiu HY, Hsiez YW: A drug identification model developed using deep learning technologies: experience of a medical center in Taiwan. BMC health services research2020, 20(1):312.
22. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O: Deep learning with convolutional neural network in radiology. Japanese journal of radiology2018, 36(4):257-272.
23. Sun Q, Jankovic MV, Bally L, Mougiakakou SG: Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network. In: 2018 14th Symposium on Neural Networks and Applications (NEUREL): 20-21 Nov. 2018 2018; 2018: 1-5.
24. Dalal V, Carmicheal J, Dhaliwal A, Jain M, Kaur S, Batra SK: Radiomics in stratification of pancreatic cystic lesions: Machine learning in action. Cancer letters2020, 469:228-237.
25. Ferreira JR, Jr., Oliveira MC, de Azevedo-Marques PM: Characterization of Pulmonary Nodules Based on Features of Margin Sharpness and Texture. Journal of digital imaging2018, 31(4):451-463.
26. Boostani R, Karimzadeh F, Nami M: A comparative review on sleep stage classification methods in patients and healthy individuals. Computer methods and programs in biomedicine2017, 140:77-91.
27. Luo L, Xiong Y, Liu Y, Sun X: Adaptive Gradient Methods with Dynamic Bound of Learning Rate. In: arXiv e-prints. 2019.
28. Zeiler MD: ADADELTA: An Adaptive Learning Rate Method. In: arXiv e-prints. 2012.
29. Hwang AY, Gums TH, Gums JG: The benefits of physician-pharmacist collaboration. The Journal of family practice2017, 66(12):E1-e8.
30. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W: Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer genomics & proteomics2018, 15(1):41-51.
31. Krell MM, Wilshusen N, Seeland A, Kim SK: Classifier transfer with data selection strategies for online support vector machine classification with class imbalance. Journal of neural engineering2017, 14(2):025003.
32. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y: Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology2017, 2(4):230-243.
33. Han H, Jiang X: Overcome support vector machine diagnosis overfitting. Cancer informatics2014, 13(Suppl 1):145-158.
34. Yanase T, Yanagita I, Muta K, Nawata H: Frailty in elderly diabetes patients. Endocrine journal2018, 65(1):1-11.

Figures
Figure 1

Structure of the clinical decision support system (CDSS), Long short time model (LSTM) and bidirectional-LSTM (Bi-LSTM) model. EHR: electronic health records.
Figure 2: Data management flowchart

Data management flowchart. R: visiting times; P: patient case numbers; Q: seasonal times; Lab. data: laboratory data.