Analysis of Influencing Factors of Community Management of Diabetes based on Adaptive-Lasso Logistic Regression Model

Wei Lin
Chengdu University of Traditional Chinese Medicine
Zejuan Huang
Chengdu University of Traditional Chinese Medicine
Zhiqing Cao
Chengdu University of Traditional Chinese Medicine
Jing Zhou
Chengdu University of Traditional Chinese Medicine
Yang Tian
Shizishan Community Health Service Centre
Jiawei Li
Chengdu University of Traditional Chinese Medicine
Mingyue Zheng (mingyue.zheng@adelaide.edu.au)
The University of Adelaide Adelaide Medical School https://orcid.org/0000-0002-7946-8911

Research article

Keywords: Community diabetes management, Adaptive-lasso logistic model, Influencing factors, Health technology assessment

Posted Date: January 17th, 2020

DOI: https://doi.org/10.21203/rs.2.21067/v1

Analysis of Influencing Factors of Community Management of Diabetes based on Adaptive-Lasso Logistic Regression Model

Wei Lin[1], Zejuan Huang[1], Zhiqing Cao[1], Jing Zhou[1], Yang Tian [2], Jiawei Li[1], Mingyue Zheng[34*,]

([1] School of Management, Chengdu University of Traditional Chinese Medicine, Chengdu, 611137, China;

[2] Shizishan Community Health Service Center, Chengdu, 610021, China;

[3] School of Health and rehabilitation, Chengdu University of Traditional Chinese Medicine, Chengdu, 611137, China; [4] Adelaide Medical School, University of Adelaide, Adelaide, SA 5005, Australia)

Corresponding author: Mingyue Zheng[34*], e-mail: mingyue.zheng@adelaide.edu.au

Abstract Background: Influencing factors of community management of diabetes are complex and controversial, and how to select the most effective multiple influencing factors requires in-depth research. This study aims to analyse multiple influencing factors by adaptive-lasso logistic regression model for the effectiveness of diabetes management of the community to improve the efficiency and reduce the burden of diabetes. **Methods:** A cross-sectional survey (N=1,127) was adopted to establish the adaptive-lasso logistic regression model of influencing factors for community management of diabetic patients based on cluster sampling data of diabetic patients in Chengdu city, China. By comparing with the full-variable logistic model and the ridge logistic model to find the advantages of the adaptive lasso-logistic regression model in community diabetes management. **Results:** A total of 1,127 diabetic patients were included in the cross-sectional survey. The latest fasting blood glucose was included in the analysis. Among the included population, 90.6% of them had a fasting glucose level higher than 6.1mol/L, and 9.4% of them were below 6.1mol/L. By cross-validation, after folding eight times, the variables involved in the Adaptive lasso-logistic regression model include age, education level, main source of income, marital status, average monthly income, free medical service, basic medical insurance for residents, hospital history, number of follow-up evaluations by family doctor team, voluntary participation in community blood glucose measurement. The AIC and BIC criteria of adaptive lasso-logistic regression model were 2062 and 1981, which were lower than the full-variable logistic model (2349, 2023) and the ridge logistic model (2312, 2013). From the perspective of time cost, the adaptive-lasso logistic regression model was better than the other two models. **Conclusions**: The adaptive-lasso logistic regression model can be used to analyse the influencing factors of community management in patients with diabetes. Community intervention and intensive management measures can significantly improve the blood glucose status of patients with diabetes.

[Key words] Community diabetes management, Adaptive-lasso logistic model, Influencing factors, Health technology assessment

1. Background

The global costs of diabetes and its consequences are large and will substantially increase more than $2.1 trillion in 180 countries by 2030 [1], especially in developing countries. Community-based diabetes management has become a cost-effective strategy for controlling and managing diabetes in primary settings. The community-based intervention positively reduced the HbA1c level of diabetic people [2, 3], improved disease control [4], and improved psychosocial outcomes [5].

In primary care and community settings, interventions targeting diabetes management were better targeted at individuals with poor glycemic control, old age, and family history [6, 7]. Moreover, younger age and lower educational attainment were associated with lower probability of meeting the goals of diabetes management [8]. However, the influencing factors of diabetic patients based on community management are complex, and how to choose the most effective influencing factors requires in-depth research.

In terms of community health service, it proposed that patients' satisfaction with community health service was moderate, high satisfaction with the community health service shown better medication adherence and regular self-monitoring of blood glucose, and these associations varied by socioeconomic status [9]. In particular, team-based care can improve patients' glucose levels, blood pressure, and lipid levels based on community preventive services task force [10]. However, what effective services the community health service or the family doctor should provide to a diabetic patient is still to be investigated (i.e. counselling, education, free health service).

The lasso model was first proposed by Tibshirani (1996) [11]. This model is to compress the parameters and make some regression coefficients gradually smaller or even close to zero. It has the advantages of subset selection and ridge regression. Huang (2008) proposed a new variable selection and estimation method based on the lasso model, which can predict the correlation pattern between variables [12]. Furthermore, many studies adopted the lasso model to screen variables and use them for optimisation and optimisation [13-15]. However, according to the author's knowledge, no adaptive-lasso logistic regression model has been used to investigate the influencing factors of diabetes community management.

Therefore, this study aimed to analyse which influencing factors may be more effective to achieve diabetes management targets in the community by the adaptive-lasso logistic regression model.

2. Methods

2.1 Data Sources

A cross-sectional survey method (N=1,127) based on the multi-stage cluster sampling was used to survey selected samples from three communities from January 2019 to August 2019. We surveyed the registered residents with the help from community health service centres in the Shizishan Community, Longzhou Road Community, and Lianxin Community of Chengdu, Sichuan Province. The inclusion criteria were through community mobilisation and publicity by the community health service centre. To avoid bias in the study design, we conducted a multidisciplinary expert demonstration and formulated reasonable inclusion and exclusion criteria to ensure higher reliability and validity when formulating the questionnaire. Random number tables were used to select registered diabetic patients in three communities in this study. The survey included demographic information (i.e. gender, age, marital status, education, number of children), self-diabetes management status (i.e. self-assessment of glycemic management, hospital history), basic situation of living (i.e. main source of income, average monthly income, medical payment methods), diabetes management in the community in the past six months (i.e. number of follow-up evaluations by the family doctor team, number of health consultations and free consultations, whether to participate in community blood glucose measurement voluntarily) and latest fasting blood glucose level and other information. Family doctor follow-up records, blood glucose records, and health consultations and free consultations were derived from the residents' health management files and the Chengdu regional health information platform.

Participants whose home address or contact details were unchanged for more than three years were included in this study. The collection questionnaire was double-entry by two individually researches to avoid bias in the data collection process. During the process of data collection and entry, no data loss occurred.

Each participant signed an informed consent form, and they volunteered to withdraw from the questionnaire process at any time without reason. For illiterate and semi-illiterate participants, data were collected by investigators reading informed consent and reading questionnaires. Data collectors were not involved in the data analysis process to avoid bias. The Chronbach 's α coefficient of the knowledge, attitude and behaviour evaluation scale for people with diabetes was 0.757, which can be considered as a good internal consistency.

2.2 Lasso Algorithm

Lasso algorithm is a regularisation method based on parameter estimation and variable selection. The parameter estimation is defined as follows:

$$\hat{\beta}_{\text{lasso}} = \arg\ \min{}^{2} \| Y - \sum_{j=1}^{p} X_j \beta_j \|^2 + \lambda \sum_{j=1}^{p} |\beta_j| \qquad (1-1)$$

In formula (1-1), λ is a regularised non-negative parameter, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^T$ is the regression coefficient, $X_j = (x_{1j}, x_{2j}, \ldots, x_{nj})^T$ and $j = 1, 2, \cdots, p$ are predictor variables, $X = (X_1, X_2, \ldots, X_p)^T$ is the predictor matrix, $Y = (Y_1, Y_2, \ldots, Y_n)^T$ is the response variable. $\lambda \sum_{j=1}^{p} |\beta_j|$ is called ‘ l penalty’, when λ increases, the lasso model allows the coefficients to approach zero; when $\lambda \to \infty$, the coefficients almost reach zero.

The improvement of the adaptive-lasso method is to add different weights to different coefficients, and its expression is defined as:

$$\hat{\beta}^{*(n)} = \arg\min \|Y - \sum_{j=1}^{p} X_j \beta_j\|^2 + \lambda_n \sum_{j=1}^{p} \hat{w}_j |\beta_j| \qquad (1\text{-}2)$$

In formula (1-2), $\hat{w}_j = \frac{1}{|\hat{\beta}_j|^{\gamma}} (\gamma > 0)$, $j = 1, 2, \cdots, p$, where $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_p)^T$ is the coefficient estimated value obtained by the ordinary least squares method. Its weight vector expression is as follows:

$$\hat{W} = (\hat{w}_1, \hat{w}_2, \cdots, \hat{w}_p)^T = \left(\frac{1}{|\hat{\beta}_1|^{\gamma}}, \frac{1}{|\hat{\beta}_2|^{\gamma}}, \cdots, \frac{1}{|\hat{\beta}_p|^{\gamma}} \right) = \frac{1}{|\hat{\beta}|^{\gamma}} \quad (\gamma > 0) \qquad (1\text{-}3)$$

2.3 Adaptive-lasso Estimation of Logistic Regression Models

Based on logistic regression, $\hat{\beta}$ is defined as $\hat{\beta} = \arg\min^2 \left(-L(\beta) + \lambda \sum_{j=1}^{p} |\beta_j| \right)$, where $L(\beta)$ is its log-likelihood function. When the fasting glucose level of included patients is less than or equal to 6.1mol/L, y_i will be defined as 0; when the fasting glucose level is greater than 6.1mol/L, y_i will be defined as 1, then:

$$p_i(y_i) = p_i^{y_i}((1-p_i)^{1-y_i}) \quad i = 1, 2, \cdots, n$$

From this, we can obtain the log-likelihood function of the joint density function for n samples as follows: $L(\beta) = \sum_{i=1}^{n} p_i y_i = \sum_{i=1}^{n} \left(y_i \log \frac{p_i}{1-p_i} + \log(1-p_i) \right)$

Based on logistic regression, the estimated value of lasso can be expressed as follows:

$$\hat{\beta}_{\text{lasso}} = -\sum_{i=1}^{n} \left(y_i \left(\beta_{ij} + \sum_{i=1}^{p} \beta_i x_i \right) - \log\left(1 + \exp\left(\sum_{i=1}^{p} \beta_i x_i \right) \right) \right) + \lambda \sum_{j=1}^{p} |\beta_j|$$

2.4 Statistical Methods

Data were entered using Epidata (Version 3.1), data analysis was performed by R (Version

3.5.3), and chi-square test and adaptive lasso-logistic regression were used to analyse the influencing factors of diabetes management. Due to the inconsistency of the dimensions, the data in this article is standardised using lar () in R, so that different eigenvalues have the same scale. Among them, the lasso-logistics regression model was completed by the glmnet package in R.

2.5 Model Analysis

From January 2019 to August 2019, the questionnaire survey was conducted among diabetic patients in community hospitals or primary health service centres in Chengdu, Sichuan Province. In terms of sample size, one independent variable of the regression needs at least ten samples to support. A total of 15 independent variables were selected in this study. Considering the replacement of some samples in the sampling, elimination, and poor questionnaire compliance will lead to a reduction in the number of evaluable cases. Therefore, a total of 1,200 questionnaires were distributed, and 1,127 valid questionnaires were collected. Participants can exit without giving any reason during the process. No participant withdrew from this study, but 73 of the surveys were considered invalid because they were incomplete. The effective response rate was 93.9%. The age range was 28-97 years old. Regarding self-assessment of glycemic management, 32.5% of patients felt satisfied, 59.8% of patients felt generally satisfied, and 6.8% of patients felt not very satisfied, and 6.9% of patients felt dissatisfied (See Table 1).

2.6 Analysis of Influencing Factors of Self-management in Patients with Diabetes

The univariate chi-square test was performed for each categorical variable included in the model and the latest fasting blood glucose (See Table 1). It can be seen that age, education level, main source of income, marital status, average monthly income, medical payment methods (basic medical insurance for residents, basic medical insurance for employee and free medical service), and hospital history were statistically significant ($p<0.05$).

Table 1 Variables and fasting glucose control

Variable		**Proportion (%)**	χ^2
Gender	Male (540)	47.90%	0.393
	Female (587)	52.10%	
Educational level	Illiterate or semi-illiterate (106)	9.40%	15.062^{**}
	primary school (395)	35.00%	
	Junior school (337）	29.90%	
	High school (251）	22.20%	
	University and above (38)	3.40%	
Age	18-44 (Youth) (38)	3.40%	14.594^{**}
	45-59 (Middle-aged people) (163)	14.50%	
	60-74 (Younger seniors) (627)	55.60%	
	75 or above (Seniors) (299)	26.50%	
Marriage status	Yes (838)	74.40%	15.304^{**}
	No (289)	25.60%	
Main source of income	Retirement salary (462)	41.00%	12.331^{**}
	Social Security (405)	35.90%	
	Salary income (154)	13.70%	
	Other (106)	9.40%	
Average monthly income	1,000 yuan or less (183)	16.20%	24.073^{***}
	1,000-2,000 yuan (443)	39.30%	
	Over 2,000 yuan (501)	44.40%	
Basic medical insurance for residents	Yes (587)	52.10%	14.274^{**}
	No (540)	47.90%	
Basic medical insurance for employee	Yes (518)	46.00%	13.461^{**}
	No (609)	54.00%	
Free medical service	Yes (39)	3.40%	16.013^{**}
	No (1,088)	96.60%	
Hospital history	Yes (241)	21.40%	13.843^{**}

	No (886)	78.60%	
Self-assessment of glycemic management	Satisfied (366)	32.50%	11.89**
	Generally satisfied (674)	59.80%	
	Not very satisfied (77)	6.80%	
	Dissatisfied (10)	0.90%	
Number of follow-up evaluations by the family doctor team	0 times (16)	1.20%	
	1 time (94)	8.3%	13.181**
	2 times and above (1,017)	90.50%	
Voluntary participation in community blood glucose measurement	Yes (1083)	96.10%	8.93**
	No (44)	3.90%	

*: $p<0.1$; **: $p<0.05$.

Notes: 1) Hospital history means as long as the patient self-reports that they have been hospitalised once, the answer is yes, otherwise the answer is no; 2) Self-assessment of glycemic management is a relatively subjective self-evaluation of patients and their satisfaction with blood glucose management; 3) The number of follow-up evaluations by the family doctor team refers to the number of family doctor follow-ups arranged by the community health service centre during the period from January 2019 to August 2018; 4) Voluntary participation in community blood glucose measurement refers to the determination of patient's blood glucose and other biochemical information regularly held by the community health service centre, and whether the patient comes voluntarily.

2.7 Adaptive-lasso Logistic Regression Analysis of Influencing Factors of Diabetes Self-management

Due to the inconsistency of the dimensions, the data in this study was standardised by using lar () in R, so that different feature values had the same scale. This study adopted the latest fasting blood glucose measurement (0 or 1) as the dependent variable. The following factors that may affect the management of diabetes were assigned as independent variables (See Table 2). The logistic regression model of adaptive-lasso variable selection was introduced [4], and the coefficients of each selected variable were estimated to analyse the factors affecting diabetes management. The Lasso-logistics regression model of this study is completed by the glmnet package in R. The relationship between the model error and was obtained through cross-validation. The number of folding times was eight times. The selected variables and parameter estimates were shown in Table 3.

Table 2 Variable assignment

Variable	**Symbol**	**Assignment**
Latest fasting blood glucose	Y	Measured value <6.1: Y = 0; Measured value> 6.1: Y = 1
Age	X_1	0-17：X_1=0；18-44：X_2=1；45-59：X_3=2；60-74：X_4=3; 75 and above：X_5=4
Educational level	X_2	Illiterate or semi-illiterate: X_2 = 0; Primary school: X_2 = 1; Junior school: X_2 = 2; High school: X_2 = 3; University and above: X_2 = 4;
Marriage status	X_3	Yes: X_3=0; No: X_3=1
Main source of income	X_4	Retirement salary: X_4 = 0; Social security: X_4 = 1; Salary income: X_4 = 2; others: X_4 = 3
Average monthly income	X_5	1000 yuan or less: X_5=0；1000-2000 yuan: X_5=1; Over 2000 yuan: X_5=2
Free medical service	X_6	Yes: X6=0; No: X6=1
Basic medical insurance for residents	X_7	Yes: X_7=0; No: X_7=1
Basic medical insurance for employee	X_8	Yes: X_8=0; No: X_8=1
Hospital history	X_9	Yes: X_9=0; No: X_9=1
Gender	X_{10}	Male: X_{10}=0; Female: X_{10}=1
Self-assessment of glycemic management	X_{11}	Satisfied: X_{11} = 0; Generally satisfied: X_{11} = 1; Not very satisfied: X_{11} = 2; Dissatisfied: X_{11} = 3
Number of follow-up evaluations by the family doctor team	X_{12}	$X_{12} \in N^+$

Number of health consultations and free consultations	X_{13}	$X_{13} \in N^{+}$
Voluntary participation in community blood glucose measurement	X_{14}	Yes：X_{14}=0; No: X_{14}=1

Notes: 1) Self-assessment of glycemic management is a relatively subjective self-evaluation of patients and their satisfaction with blood glucose management; 2) The number of follow-up evaluations by the family doctor team refers to the number of family doctor follow-ups arranged by the community health service centre during the period from January 2019 to August 2018; 3) The number of health consultations and free consultations refers to the number of times patients participated in health consultations and free consultations organised by community health service centres; 4) Voluntary participation in community blood glucose measurement refers to the determination of patient's blood glucose and other biochemical information regularly held by the community health service centre, and whether the patient comes voluntarily.

Table 3 Parameter estimation

Variable	**Adaptive-lasso logistic**	**Full-variable logistic**	**Ridge logistic**
Intercept	-2.092	-2.797	-2.334
Age			
18-44 years old	0.155	0.139	0.094
45-59 years old	0.683	0.703	0.625
60-74 years old	0.654	0.701	0.601
Educational level			
Illiterate or semi-literate	0.000	-0.909	-0.423
primary school	0.000	0.549	0.434
Junior school	0.139	0.119	0.087
High school	0.748	0.779	0.754
Marital status	-0.412	-0.438	-0.432
Main source of income			
Retirement salary	-0.801	-0.824	-0.820
Social Security	-0.991	-0.995	-0.993
Salary income	0.000	-0.540	0.000

Average monthly income			
1000 yuan or less	0.000	-0.704	0.384
1000-2000 yuan	0.073	0.068	0.067
Medical payment methods			
Free medical service	0.427	0.206	0.452
Basic medical insurance for residents	0.637	0.852	0.689
Basic medical insurance for employee	0.000	0.689	0.000
Hospital history	0.584	1.489	0.761
Number of follow-up evaluations by the family doctor team	0.737	0.724	0.692
Number of health consultations and free consultations	1.049	0.933	0.847
Latest fasting blood glucose	0.000	0.432	0.323
BIC	2062	2349	2312
AIC	1981	2023	2013

Notes: 1) Hospital history means as long as the patient self-reports that they have been hospitalised once, the answer is yes, otherwise the answer is no; 2) Medical payment methods include Basic medical insurance for residents, Basic medical insurance for employee, and free medical service; 3) Self-assessment of glycemic management is a relatively subjective self-evaluation of patients and their satisfaction with blood glucose management; 4) The number of follow-up evaluations by the family doctor team refers to the number of family doctor follow-ups arranged by the community health service centre during the period from January 2019 to August 2018.

Comparing the three models from Table 3, it was found that the adaptive-lasso logistic model was more significant in selecting variables, and the model was more concise. The Adaptive-lasso logistic model had the smallest AIC and BIC criteria, and the full-variable logistic model had the largest AIC and BIC criteria.

2.8 Comparison of Model Prediction Accuracy

To predict the results of diabetes management, the established full-variable logistic model, adaptive-lasso logistic model and ridge logistic model were used. Table 4 shown the prediction accuracy of the models with different proportions of the training set. The prediction accuracy of the full-variable logistic model and the adaptive-lasso logistic model was slightly higher than that of the ridge logistic model (See Table 4). However, the prediction accuracy of the three models was not significantly different, and the reason for the small difference in accuracy might relate to data collection. But only from the time required to collect data, the running time of the adaptive-lasso logistic model was less than the prediction time of the full-variable logistic model. When the prediction accuracy of the three models was close, or there was not much difference, from the perspective of time cost, the adaptive-lasso logistic model was better than the other two models.

Table 4. Prediction accuracy of the models

	Proportion of training set		
Model category	**95%**	**90%**	**80%**
Adaptive-lasso logistic model	61.34%	61.89%	62.76%
Full-variable logistic model	61.52%	62.01%	62.50%
Ridge logistic model	60.21%	61.48%	62.18%

3. Results and Discussion

3.1 Community management of people with diabetes affected by multiple factors

Community diabetes management in the are complex and are affected by many variables. Adaptive-lasso logistic model, full-variable logistic model and ridge logistic model were used to screen variables, and then age, education, marital status, main source of income, monthly average income, basic medical insurance for employees, and whether they have been hospitalised were all included in the above three models. It indicated that these factors were related to self-management factors in patients with diabetes. Our findings indicated that the effects of age, marital status, and hospital history on self-management of patients with diabetes were significant, this result was consistent with Lian (2019) and Bansal (2018) [16, 17]. It also found that the younger the diabetic patients are, the more rigorous they are. As age increases, the elderly need to strengthen their self-management of diabetes [13, 16]. Besides, educational level [8, 14] was an influential factor in community management of diabetes, and the higher the educational level, the better the self-glycemic control. In terms of public medical care, we also found that compared with those who did not receive public medical care, those who had public medical care had a slight advantage in blood glucose management. Besides, patients with a history of hospitalisation were better at controlling blood glucose than patients without a history of hospitalisation.

3.2 Adaptive-lasso logistic model can be used to analyse the influencing factors of community management in diabetic patients

The parameter estimation of the adaptive-lasso logistic model was both unbiased and has the advantages of ridge regression and subset selection [13]. In this study, the adaptive-lasso algorithm was introduced into a logistic regression model to analyse the influencing factors among people with diabetes and to evaluate community diabetes management. According to the authors' knowledge, using this model to discuss the influencing factors of blood glucose management from the perspective of community diabetes management has not been reported. The adaptive-lasso algorithm was used in the logistic regression model to achieve the purpose of filtering variables and simplifying the model in this study. Based on the cross-sectional survey data, comparing the fitted full-variable logistic model and the ridge logistic model, the AIC and BIC criteria of the adaptive-lasso logistic model were the smallest; when making predictions, the time cost of the adaptive-lasso logistic model was the least. Through analysis and comparison, the adaptive-lasso logistic model was more concise than the variables selected by full-variable logistic, and the model was more compressed. The adaptive-lasso logistic model considered the variables that affect diabetes self-management included age, education level, the main source of income, marital status, average monthly income, free medical service, basic medical insurance for residents, hospital history, number of follow-up evaluations by family doctor team, voluntary participation in community blood glucose measurement.

Community intervention and intensive management measures can significantly improve blood glucose management among patients with diabetes. In particular, increasing the number of follow-up evaluations, health consultations and free consultations by the family doctor team will significantly affect the blood glucose management effect of diabetic patients, our results were consistent with the results of Aponte (2017) [4]. It seems that practical strategies to improve diabetes self-management based on the community include social support for family members, health cares, and community members, and local free or low-cost diabetes education materials and courses [18]. However, this study also had some limitations. The samples were mainly from three communities, the sample size was relatively small, and the prediction accuracy of the three models was not much different, which may be related to the insufficient sample size. Moreover, the survey of this study could not include lifestyle influencing factors such as physical activity and diet, considering the target population was based on the community setting.

4. Conclusions

The adaptive-lasso logistic model can be used in the analysis of community management factors for diabetes. It can accurately screen out the factors that affect diabetes management. The obtained model can better explain the indicators of these influencing factors and provide advice for primary care settings and communities. Diabetes management is affected by many factors, and relevant knowledge and education should be strengthened for elderly, non-married, low-educated, low-income people, diabetic patients without public medical care and no history of hospitalisation. Moreover, the number of follow-up evaluations of family doctor teams and the frequency of health consultations and free consultations should be increased to achieve optimal management of diabetes based on community.

5. Declarations

Ethics approval and consent to participate

This study was an observational study obtained consent from the ethics committee of Chengdu University of Traditional Chinese Medicine. All participants signed an informed consent form, and they volunteered to withdraw from the questionnaire process at any time without reason. For illiterate and semi-illiterate participants, data were collected by investigators reading informed consent and reading questionnaires. Data collectors were not involved in the data analysis process to avoid bias.

Consent for publication

All participants agreed that we use the collected data for academic publications.

Availability of data and materials

All data analysed during this study are included in this manuscript.

Competing Interests

All authors of the study stated that there were no conflicts of interest.

Funding

This work was supported by the Primary Health Care Development Research Center of Sichuan Province (grant number SWFZ17-Y-33) and Chengdu University of Traditional Chinese Medicine (grant number ZRQN2019010) hold by Wei Lin. Chengdu University of Traditional Chinese Medicine (grant number ZRQN2019009) hold by Mingyue Zheng.

Authors' contributions

Wei Lin and Mingyue Zheng conceptualised the study, drafted the initial manuscript, reviewed and revised the manuscript. Jing Zhou, Zejuan Huang and Zhiqing Cao, Yang Tian collected data, carried out the initial analyses. Wei Lin and Mingyue Zheng mainly discussed and edited the manuscript. Wei Lin, Mingyue Zheng, Jiawei Li designed the data collection instruments, coordinated and supervised data collection, and critically reviewed the manuscript. All authors approved the final manuscript as submitted and agree to be responsible for all aspects of the work.

Acknowledgements

We are very grateful for the support from the doctors and nurses of the community health service centre during the questionnaire distribution and data collection process from the Lion Mountain Community, Longzhou Road Community, and Lianxin Community in Chengdu, Sichuan Province.

Abbreviations

AIC: Akaike information criterion

BIC: Bayesian Information Criterion

References

1. Bommer, C., et al., *Global Economic Burden of Diabetes in Adults: Projections From 2015 to 2030.* Diabetes Care, 2018. **41**(5): p. 963-970.
2. Pamungkas, R.A. and K. Chamroonsawasdi, *HbA1c reduction and weight-loss outcomes: a systematic review and meta-analysis of community-based intervention trials among patients with type 2 diabetes mellitus.* International Journal of Diabetes in Developing Countries, 2019. **39**(2): p. 394-407.
3. Little, T.V., et al., *Community health worker interventions for Latinos with type 2 diabetes: a systematic review of randomized controlled trials.* Curr Diab Rep, 2014. **14**(12): p. 558.
4. Aponte, J., et al., *Health effectiveness of community health workers as a diabetes self-management intervention.* Diab Vasc Dis Res, 2017. **14**(4): p. 316-326.
5. Davies, M.J., et al., *A community based primary prevention programme for type 2 diabetes integrating identification and lifestyle intervention for prevention: the Let's Prevent Diabetes cluster randomised controlled trial.* Prev Med, 2016. **84**: p. 48-56.

6. Murphy, M.E., et al., *Improving risk factor management for patients with poorly controlled type 2 diabetes: a systematic review of healthcare interventions in primary care and community settings.* BMJ Open, 2017. **7**(8): p. e015135.

7. Omar, S.M., et al., *Prevalence, risk factors, and glycaemic control of type 2 diabetes mellitus in eastern Sudan: a community-based study.* Ther Adv Endocrinol Metab, 2019. **10**: p. 2042018819860071.

8. Siegel, K.R., et al., *Prevalence of Major Behavioral Risk Factors for Type 2 Diabetes.* Diabetes Care, 2018. **41**(5): p. 1032-1039.

9. Yin, T., et al., *Socioeconomic status moderates the association between patient satisfaction with community health service and self-management behaviors in patients with type 2 diabetes: A cross-sectional survey in China.* Medicine (Baltimore), 2019. **98**(22): p. e15849.

10. Community Preventive Services Task Force. Electronic address, y.c.g., *Team-Based Care to Improve Type 2 Diabetes Management: Recommendation of the Community Preventive Services Task Force.* Am J Prev Med, 2019. **57**(1): p. e27-e29.

11. Tibshirani R. *Regression shrinkage and selection via the lasso.* Journal of the Royal Statistical Society: Series B (Methodological). 1996;**58**(1):267-88.

12. Jian Huang, S. Ma, and C. Zhang, *Adaptive Lasso for Sparse High-dimensional Regression Midels.* Statistica Sinica 2008. **18**, 1603-1618.

13. Sun M, Tian M. *A Class of Derivative-Free CG Projection Methods for Nonsmooth Equations with an Application to the LASSO Problem*. Bulletin of the Iranian Mathematical Society. 2019:1-23.

14. Shukor, S., et al., *Quantitative assessment of LASSO probe assembly and long-read multiplexed cloning.* BMC Biotechnol, 2019. **19**(1): p. 50.

15. Jung, A. and N. Tran, *Localized Linear Regression in Networked Data.* 2019.

16. Lian, J., et al., *Long-term cost-effectiveness of a Patient Empowerment Programme for type 2 diabetes mellitus in primary care.* Diabetes Obes Metab, 2019. **21**(1): p. 73-83.
17. Bansal, V., et al., *Inpatient diabetes management by specialized diabetes team versus primary service team in non-critical care units: impact on 30-day readmission rate and hospital cost.* BMJ Open Diabetes Res Care, 2018. **6**(1): p. e000460.
18. Purnell, T.S., et al., *Perceived Barriers and Potential Strategies to Improve Self-Management Among Adults with Type 2 Diabetes: A Community-Engaged Research Approach.* Patient, 2016. **9**(4): p. 349-58.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- STROBEchecklistcrosssectionalWeiLin.doc
- InformedConsentforDiabetesCommunityQuestionnaire.docx