Research Article

The Research of Improved Grey GM (1, 1) Model to Predict the Postprandial Glucose in Type 2 Diabetes

Yannian Wang, 1 Fenfen Wei, 1 Changqing Sun, 2 and Quanzhong Li 3

1 School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China
2 College of Public Health, Zhengzhou University, Zhengzhou 450001, China
3 Department of Endocrinology, Henan People’s Hospital, Zhengzhou 450003, China

Correspondence should be addressed to Quanzhong Li; liquanzhongzzu@gmail.com

Received 4 February 2016; Revised 12 April 2016; Accepted 27 April 2016

Academic Editor: Yoshifumi Saisho

Copyright © 2016 Yannian Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Diabetes may result in some complications and increase the risk of many serious health problems. The purpose of clinical treatment is to carefully manage the blood glucose concentration. If the blood glucose concentration is predicted, treatments can be taken in advance to reduce the harm to patients. For this purpose, an improved grey GM (1, 1) model is applied to predict blood glucose with a small amount of data, and especially in terms of improved smoothness it can get higher prediction accuracy. The original data of blood glucose of type 2 diabetes is acquired by CGMS. Then the prediction model is established. Finally, 50 cases of blood glucose from the Henan Province People’s Hospital are predicted in 5, 10, 15, 20, 25, and 30 minutes, respectively, in advance to verify the prediction model. The prediction result of blood glucose is evaluated by the EGA, MSE, and MAE. Particularly, the prediction results of postprandial blood glucose are presented and analyzed. The result shows that the improved grey GM (1, 1) model has excellent performance in postprandial blood glucose prediction.

1. Introduction

As one of the major diseases that harm human health, diabetes can lead to many complications, including atherosclerosis, blindness, renal failure, and feet disorders with risk of amputation [1, 2]. According to the WHO, there are nearly 347 million diabetics in the world. The clinical presentation of diabetes includes hyperglycemia and hypoglycemia. And the main purpose of the clinical treatment of diabetes is to keep good control of blood glucose concentration. If the future blood glucose concentration is predicted, doctors and patients can take some action in advance to reduce the harm to patients. Compared with fasting glucose, postprandial blood glucose is more harmful because it tends to fluctuate. The Guideline for Management of Postprandial Blood Glucose also emphasizes that controlling of postprandial blood glucose has extremely profound significance. In order to reduce the risk of complications [3] and carefully control blood glucose levels in advance, effective postprandial blood glucose prediction method should be studied and it can provide support for doctors and patients.

CGMS (Continuous Glucose Monitoring System) is a device that is placed on the patient and used to measure patient’s blood glucose every 5 minutes. Based on the blood glucose data provided by CGMS, many kinds of prediction methods of blood glucose were proposed, such as adaptive blood glucose prediction model [4], AR (autoregressive) model [5], neural network prediction model [6], and SVM (support vector machine) model [7]. Peng et al. [4] applied Kalman filter to smooth the blood glucose data from the CGMS, using AR model to build up the blood glucose prediction model, and the result showed that the blood glucose changes can be dynamically captured and the future blood glucose can be predicted. Wang and An [5] also adopted the AR model in predicting blood glucose; the result showed that the prediction was accurate with simple calculation, but their research did not take into account the smoothness of the original data. Tresp et al. [6] utilized neural network algorithm to predict blood glucose and found that their model had good tracking ability. Georga et al. [7] employed SVM algorithm to predict blood glucose and found that it had good prediction effectiveness when it had a large amount of
data. In terms of selecting or training model parameters, if the original data sequence is longer, their prediction results are reliable, but if the original data sequence is shorter, due to inadequate information and lack of significant regularity, their forecasting accuracy is low.

The grey GM (1, 1) model not only has simple principle, less samples, easy calculation, high forecasting accuracy, and easy inspection but also can preprocess the original data, obtain better smoothness, and predict more effectively. In this paper, an improved grey GM (1, 1) model was proposed to make predictions in 30 min in advance on 72 hours of blood glucose and 2 hours of postprandial blood glucose, respectively; the experiment was performed on the MATLAB and the results were compared with AR model.

2. Materials and Methods

Grey GM (1, 1) model has been widely applied in many fields, such as economy, science, and education [8]. Grey GM (1, 1) model is a kind of homogeneous exponential growth model based on the accumulation generation sequence and the least squares method. The growth trend of original data has great influence on the accuracy of prediction. If the original data sequence is smooth, the more close to the exponential growth it is, the higher prediction precision the model can produce. The improved grey prediction model preprocesses the original data to improve the smoothness of the data sequence and greatly improve the prediction precision and the predictive value of original data sequence was obtained through an inverse transformation.

2.1. The Blood Glucose Prediction Model Based on the Improved Grey GM (1, 1) Model

Suppose an original data sequence \(X^{(0)}\) is as the following formula:

\[ X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)), \]  

where \(x^{(0)}(i) > 0, i = 1, 2, \ldots, n.\) The steps of establishing the blood glucose prediction model based on the data sequence \(X^{(0)}\) are as follows.

(1) As a logarithmic transformation on \(X^{(0)},\) mark \(y^{(0)}(i) = \ln x^{(0)}(i), i = 1, 2, \ldots, n.\) The data sequence \(Y^{(0)}\) can be processed as the following formula:

\[ Y^{(0)} = (y^{(0)}(1), y^{(0)}(2), \ldots, y^{(0)}(n)). \]  

(2) Generate the accumulation generation sequence \(Y^{(1)}\) by 1-AGO, as in the formula

\[ Y^{(1)} = (y^{(1)}(1), y^{(1)}(2), \ldots, y^{(1)}(n)), \]  

where \(y^{(1)}(1) = y^{(0)}(1), y^{(1)}(k) = \sum_{i=1}^{k} y^{(0)}(i) (k = 2, 3, \ldots, n).\)

(3) Through the first-order accumulative generation sequence \(Y^{(1)},\) GM (1, 1) model is established; a first-order differential equation can be gotten as the formula

\[ \frac{dy^{(1)}(t)}{dt} + ay^{(1)}(t) = b, \]  

where \(a\) is the development coefficient and \(b\) is the control variable. Get the corresponding form of grey differential equation as the following formula:

\[ x^{(0)}(i) + ax^{(1)}(i) = b, \ i = 2, 3, \ldots, \]  

(5)

(4) To solve the parameters \(a\) and \(b,\) parameters \(\theta = [a, b]^T\) can be determined by the least square method as the formula

\[ \theta = [B^T B]^{-1} B^T Y, \]  

where

\[ B = \begin{bmatrix}
-z^{(1)}(2) & 1 \\
-z^{(1)}(3) & 1 \\
\vdots & \vdots \\
-z^{(1)}(n) & 1
\end{bmatrix}, \]  

\[ z^{(1)}(k) = \frac{1}{2} \left[ y^{(1)}(i) + y^{(1)}(i - 1) \right], \]  

\[ Y = \left( x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n) \right)^T. \]  

(6) The known initial condition is \(\vec{y}^{(1)}(1) = y^{(1)}(1) = y^{(0)}(1),\) put it into formula (6) and formula (5) to obtain the generated data sequence \(\vec{y}^{(1)}\) as the following formula:

\[ \vec{y}^{(1)}(i) = \left( y^{(0)}(1) - \frac{\vec{b}}{\vec{a}} \right) e^{-\frac{a(i-1)}{a}} + \frac{\vec{b}}{a} \]  

\[ (i = 2, 3, \ldots, n). \]  

(7) The known inverse transformation is as formula

\[ \vec{x}^{(0)}(i) = e^{\vec{y}^{(0)}(i)}. \]  

(11)

Put (10) into (11) and get the forecast sequence of the original blood data sequence \(\vec{x}^{(0)}\) as the formula

\[ \vec{x}^{(0)}(i) = e^{-1}(\vec{x}^{(0)}(1) - \vec{b}/\vec{a} e^{\vec{y}^{(0)}(1)}), \ i = 2, 3, \ldots, n. \]  

(12)

Put \(i = 2, 3, \ldots, n\) into formula (12), the fitted values of original data can be gotten. When \(i > n,\) the predictive values of blood glucose can be gotten.
2.2. The Ways of Testing the Accuracy of the GM (1, 1) Model. The improved GM (1, 1) model must be strictly examined and meet some requirements before being predicted. Three ways were used to estimate the accuracy of the GM (1, 1) model: relative error size test, posterior deviation test, and correlation test.

(1) The Relative Error Size Test. The relative error size test is a kind of arithmetic test by intuitively comparing the data point by point, which observes whether the relative error meets the requirement by comparing the prediction data and the actual data of blood glucose. The relative error of GM (1, 1) model is \( \varepsilon(i) \) as formula

\[
\varepsilon(i) = \frac{e(i)}{x^{(0)}(i)} \times 100\% = \frac{x(i) - \hat{x}(i)}{x^{(0)}(i)} \times 100\% ,
\]

where \( e(i) = x^{(0)}(i) - \hat{x}^{(0)}(i) \), \( i = 1, 2, \ldots, n \). The average relative error of GM (1, 1) model is \( \bar{\varepsilon} \) as the following formula:

\[
\bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^{n} |\varepsilon(i)| .
\]

The accuracy of GM (1, 1) model is \( p^* \) as the formula

\[
p^* = (1 - \bar{\varepsilon}) \times 100\% .
\]

General requirement is \( p^* > 80\% \).

(2) The Posterior Deviation Test. The posterior deviation test is of statistical method and makes inspection according to formula (2). The posterior variance ratio \( C \) and the posterior probability \( p \) are the two key indicators. The \( C \) should be as small as possible and \( p \) should be as large as possible. According to the size of the \( C \) and \( p \), the precision of the model can be divided into four levels “superior, qualified, marginal, and disqualified” \[9\]. The level of model accuracy is \( \max\{p, C\} \).

Calculate the posterior variance ratio \( C \) as formula

\[
C = \frac{S_2}{S_1},
\]

where \( S_1^2 = \frac{1}{n} \sum_{i=1}^{n} (x^{(0)}(i) - \hat{x}^{(0)})^2 \) and \( S_2^2 = \frac{1}{n} \sum_{i=1}^{n} (e(i) - \bar{\varepsilon})^2 \).

Calculate the posterior probability \( p \) as formula

\[
p = \frac{p}{|e(i) - \bar{\varepsilon}| < 0.6745S_1},
\]

where \( \bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^{n} e(i) \).

(3) The Correlation Test. The correlation test is a kind of geometry inspection, used to investigate similarity of the predictive value curve and the actual value curve. In general, the closer the geometry is, the closer the change trend is and the greater the correlation is.

The correlation degree of the GM (1, 1) model is \( \xi \) as the formula

\[
\xi = \frac{1}{n} \sum_{i=1}^{n} \xi_i,
\]

where \( \xi_i \) is the correlation coefficient of \( X^{(0)} \) and \( x^{(0)}(i) \) as the following formula:

\[
\xi_i = \min_{0 \leq j \leq n} |x^{(0)}(i) - \hat{x}^{(0)}(i)| + \rho \max_{0 \leq j \leq n} |x^{(0)}(i) - \hat{x}^{(0)}(i)| ,
\]

where \( \rho \in [0, 1] \) is the resolution ratio, \( \rho = 0.5 \) \[10\].

General requirement is \( \xi > 0.6 \) \[10\], and the greater the correlation is, the better the prediction is.

2.3. Model Evaluation Indicators. In order to evaluate forecast results, the present study applied three performance indicators: EGA (Clark Error Grid Analysis), MSE (Mean Square Error), and MAE (Mean Absolute Error).

The EGA \[11,12\] is divided into A, B, C, D, and E regions. In region A, the prediction effect is the best, while in region E it is poor. The EGA is used to evaluate the accuracy and the precision of prediction method of blood glucose and provide guidance to doctors and patients. So far, the EGA has been accepted as one of the “gold standards” in evaluating the accuracy of predicting blood glucose.

The MSE is as the formula

\[
\text{MSE} = \sqrt{\frac{\sum (e(i))^2}{n}}, \quad i = 1, \ldots, n.
\]

The MAE is as the formula

\[
\text{MAE} = \frac{\sum |e(i)|}{n}, \quad i = 1, \ldots, n.
\]

3. Results.

3.1. The Data. The data used in this study is obtained from the Henan Provincial People’s Hospital, where nurses used CGMS to acquire the blood glucose in patients as original data sequence of experiment. CGMS receives a current signal every 10 seconds and averages and converts it into blood glucose values every 5 minutes. The CGMS can continuously work at least 72 hours and store 864 blood glucose values. One of the patients’ blood glucose concentration of 2 hours after meal is shown in Figure 1(a), and one of the patients’ blood glucose concentration of 72 hours is shown in Figure 1(b). 50 cases of blood glucose will be employed to verify the prediction model based on the GM (1, 1) model and compared with the AR model.

3.2. Prediction Result of Figure 1(b). The front 200 samples’ data of Figure 1(b) is employed to establish the GM (1, 1) model for forecasting the following 664 blood glucose levels. Firstly, put the 200 samples \( n = 200 \) into formula (1) as original data sequence \( X^{(0)} \). Secondly, calculate as above the methods of the blood glucose prediction model. Then, test the accuracy of the model and do experiment which was conducted based on MATLAB. Finally, show the continuous prediction results in Figure 1(c). The prediction results of AR model are shown in Figure 1(d). By calculating, the result of
MSE of Figure 1(c) is 1.0306 and MAE is 0.3307. The result of MSE of Figure 1(d) is 1.4356 and MAE is 0.5379. The results of EGA of Figure 1(c) are shown in Figure 2, and it is clear that the majority of the error points (91.72%) lie in zone A and the remaining 8.22% in zone B.

3.3. Prediction Result of Figure 1(a). The front 18 samples data of Figure 1(a) are employed to establish the GM (1, 1) model for forecasting the following 6 blood glucose levels. Firstly, put the 18 samples \((n = 18)\) into formula (1) as original data sequence \(X(0)\). Secondly, by calculating as above the methods of blood glucose prediction model, we rewrote formula (12) of blood glucose prediction model as the following formula:

\[
\hat{x}^{(0)}(i) = 90.701e^{0.02(i-1)} - 88.7 \quad (i = 2, 3, \ldots, 18).
\]

(22)

Then, to test the accuracy of formula (22), experiment was conducted based on MATLAB. The results are \(p^* = 97.5\%, C = 0.17\), \(p = 1\), and \(\xi = 0.92\).

Finally, formula (22) of prediction model based on GM (1, 1) fulfilled the requirements; the prediction values and the errors of blood glucose are shown in Table 1. That is,

\[
\hat{x}^{(0)}(19) = 11.9, \hat{x}^{(0)}(20) = 12.4, \hat{x}^{(0)}(21) = 12.9, \hat{x}^{(0)}(22) = 13.4, \hat{x}^{(0)}(23) = 14.0, \text{ and } \hat{x}^{(0)}(24) = 14.5.
\]

As shown in Figure 2, the improved grey GM (1, 1) model can be used for predicting exactly blood glucose. According to Table 1, the postprandial blood glucose prediction result of patient A, whether in MSE or MAE, is more ideal than the continuous 72 hrs predictions.

<table>
<thead>
<tr>
<th>Sample points</th>
<th>The real value</th>
<th>The prediction value</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>12.9</td>
<td>11.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>13.1</td>
<td>12.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>12.9</td>
<td>12.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>12.7</td>
<td>13.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>12.5</td>
<td>14.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>12.4</td>
<td>14.5</td>
<td>0.9786</td>
<td>0.1631</td>
</tr>
</tbody>
</table>

MSE: mean square error; MAE: mean absolute error; the real value (mmol/L); the prediction value (mmol/L).
Table 2: The statistical errors of 50 cases.

<table>
<thead>
<tr>
<th></th>
<th>2 hrs after meal error</th>
<th>72 hrs error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
<td>MAE</td>
</tr>
<tr>
<td>AR model</td>
<td>1.8294</td>
<td>0.6549</td>
</tr>
<tr>
<td>GM (1, 1) model</td>
<td>0.9786</td>
<td>0.1631</td>
</tr>
</tbody>
</table>

MSE: mean square error; MAE: mean absolute error.

Table 3: The statistical errors in stages of 50 cases.

<table>
<thead>
<tr>
<th></th>
<th>3 hrs after meal error</th>
<th>4 hrs after meal error</th>
<th>6 hrs at night error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
<td>MAE</td>
<td>MSE</td>
</tr>
<tr>
<td>AR model</td>
<td>2.7709</td>
<td>1.4644</td>
<td>3.6996</td>
</tr>
<tr>
<td>GM (1, 1) model</td>
<td>1.1558</td>
<td>0.7697</td>
<td>2.149</td>
</tr>
</tbody>
</table>

MSE: mean square error; MAE: mean absolute error.

3.4. The Prediction Result of 50 Patients. We repeated experiments with 50 cases’ data of the blood glucose like patient A. The statistical errors of 50 cases are shown in Table 2.

As shown in Table 2, an improved grey GM (1, 1) model is applied to predict blood glucose with a small amount of data, and in particular in terms of improved smoothness it can get higher prediction accuracy than AR model, so is the prediction result of the 2 hrs after meal. This means that the improved grey GM (1, 1) model performs well not only 72 hrs of continuous prediction but also 2 hrs after the meal. In order to obtain more accurate prediction of postprandial blood glucose and provide support to doctors and patients, the improved grey GM (1, 1) model is simpler and more reliable in comparison with the AR prediction model and it only needs the historical data provided by CGMS [13, 14].

4. Discussion

In this paper, the improved grey prediction model in terms of smoothness is used to predict blood glucose of type 2 diabetes patients and the blood glucose for nearly 72 hrs and 2 hrs after meal, respectively, is predicted. The prediction result was compared with AR model. The results showed that the improved grey GM (1, 1) model has outperformed the AR model in predicting blood glucose, especially 2 hrs after meal.

In order to value the error in every stage clearly, a series of experiments was performed at three stages: 3 hrs after meal, 4 hrs after meal, and 6 hrs at night. The statistical errors in three stages of 50 cases are shown in Table 3. The error of 6 hrs at night is bigger than the error of 4 hrs after meal, and the error of 4 hrs after meal is bigger than the error of 3 hrs after meal. Additionally, as shown in Table 3, the error of 3 hrs after meal is bigger than the error of 2 hrs after meal. But the difference between the error of 6 hrs at night and the error of 2 hrs after meal is obviously bigger than the difference between the error of other stages and the error of 2 hrs after meal. This means that the prediction result of the improved grey GM (1, 1) model is worse than the AR model in the stage of 6 hrs at night. That is, the GM (1, 1) model has a more accurate predication result for the original sequences with wide fluctuation range than the ones with small fluctuation range. This disadvantage will be improved in the future research. Compared with the linear AR model, the GM (1, 1) model is better especially when the blood glucose levels suddenly rise or drop. Because the grey GM (1, 1) model is a kind of homogeneous exponential growth model, its predicting values may show the error of exponential at some sample points, but it can be resolved by correcting predicting value using real value in the GM (1, 1) model.

In a word, the method for predicting postprandial blood glucose based on the improved gray prediction model can be developed to help doctors and patients and in the future may help to develop an artificial pancreas, which adapts to
the future changing of patient’s blood glucose levels and gives better insulin outcomes as a result.

**Disclosure**

The authors alone are responsible for the content and writing of the paper.

**Competing Interests**

The authors report no conflict of interests.

**Acknowledgments**

Thanks are due to the Zhengzhou Science and Technology Research Program (Grant no. 131PPTGG409-8), the Henan Medical Science and Technology Research Program (Grant no. 201403009), and the Henan Science and Technology Research Program (Grant nos. 152102210250 and 162102310600) for financial support.

**References**


