

GM(1, 1)模型对不同范围血糖的预测性能分析

卢相月¹, 王延年², 李全忠¹

(1. 郑州大学人民医院/河南省人民医院 内分泌科, 河南 郑州, 450003;

2. 郑州大学 信息工程学院, 河南 郑州, 450001)

摘要: 目的 评价 GM(1, 1)模型在血糖预测中的性能, 分析该模型对不同范围血糖数据的预测能力。方法 以 50 例 2 型糖尿病 (T2DM) 患者的连续 72 h 血糖序列为研究对象, 基于新陈代谢算法建立 GM(1, 1)模型, 预测 5、15、30 min 后的血糖。将 50 例患者随机分成对照组 20 例和实验组 30 例, 观察分析对照组数据得到预测误差较低时的血糖范围(a~b), 以此范围将实验组数据分为 A 组(<a)、B 组(a~b)、C 组(>b), 采用 Pearson 相关性分析、非参数检验进行分析。结果 对照组筛选出预测时长为 5、15、30 min 时的血糖范围分别为 3.4~11.5、3.3~11.4、3.2~11.4 mmol/L。相关性分析显示, 不同预测时长时, B 组的预测拟合度均最好($P < 0.01$)。非参数检验表明, 同一预测时长的组间预测误差比较, 差异有统计学意义($P < 0.01$), 其中 A 组的预测误差最小, B 组次之, C 组最大。结论 GM(1, 1)模型对 3.4~11.4 mmol/L 的血糖数据预测效果更好。

关键词: GM(1, 1)模型; 2 型糖尿病; 血糖预测; 血糖范围

中图分类号: R 587.1; R 446.11 文献标志码: A 文章编号: 1672-2353(2021)09-023-06 DOI: 10.7619/jcmp.20210757

Analysis in performance of GM(1, 1) model in predicting blood glucose at different ranges

LU Xiangyue¹, WANG Yannian², LI Quanzhong¹

(1. Department of Endocrinology, Zhengzhou University People's Hospital, Henan Provincial

People's Hospital, Zhengzhou, Henan, 450003; 2. School of Information

Engineering, Zhengzhou University, Zhengzhou, Henan, 450001)

Abstract: Objective To evaluate performance of GM(1, 1) model in predicting blood glucose, and the predictive ability of this model for prediction of different blood glucose ranges. **Methods** Seventy-two-hour blood glucose sequence of fifty patients with T2DM was selected as study objects. GM(1, 1) model was established based on the metabolic algorithm to predict blood glucose levels after 5, 15 and 30 min. Fifty patients were randomly divided into control group ($n = 20$) and experimental group ($n = 30$). The blood glucose range (ranging from a to b) when the prediction error was low in the control group was obtained. According to the blood glucose range of the control group, the data of the experimental group was divided into group A (<a), group B (a to b), and group C (>b). Analysis was performed by Pearson correlation analysis and nonparametric test. **Results** The blood glucose ranges for prediction duration of 5 min, 15 min and 30 min in the control group were 3.4 to 11.5 mmol/L, 3.3 to 11.4 mmol/L and 3.2 to 11.4 mmol/L, respectively. Correlation analysis showed that group B at differed prediction durations had the best predictive fit ($P < 0.01$). Nonparametric tests indicated that there was a statistical difference in prediction error between groups at the same prediction duration ($P < 0.01$), and the group A had the minimum error, followed by group B, and group C had the maximum error. **Conclusion** GM(1, 1) model has better predictive efficacy at blood glucose range of 3.4 to 11.4 mmol/L.

Key words: GM(1, 1) model; type 2 diabetes mellitus; prediction of blood glucose; range of blood glucose

收稿日期: 2021-02-24

基金项目: 河南省科技攻关计划项目(162102310605)

通信作者: 李全忠, E-mail: loicdf@163.com

2 型糖尿病 (T2DM) 是糖尿病的主要类型, 临床发病率较高^[1-2]。T2DM 引起的慢性并发症不仅会对患者的预后及生活质量造成严重影响, 而且会给社会带来巨大的经济负担^[3]。为了更好地控制 T2DM 病情, 众多学者^[4-7] 提出预测血糖。邓聚龙教授^[8] 提出了 GM(1, 1) 模型, 该模型通过对处于白色信息 (完全已知) 和黑色信息 (完全未知) 之间的灰色信息 (部分已知、部分未知) 建立灰色微分方程来揭示其变化规律和未来趋势, 在小样本、信息获取量少的数据分析中具有优势。本研究评价了 GM(1, 1) 模型在血糖预测中的性能, 分析该模型对不同范围血糖数据的预测能力, 现报告如下。

1 资料与方法

1.1 研究资料

收集 2018 年 1—2 月在河南省人民医院内分泌科住院的 50 例 T2DM 患者作为研究对象, 其中男 28 例, 女 22 例。纳入标准: 符合 1999 年世界卫生组织 T2DM 诊断标准者; 佩戴美国美敦力公司 (Medtronic, Inc) 生产的动态血糖监测系统 (CGMS) 者。排除标准: 佩戴 CGMS 未 72 h 者或实测血糖序列有断点者; 过敏体质或有胶带状过敏史者; 妊娠期糖尿病患者; 传染性疾病患者; 有严重糖尿病急慢性并发症者或病情危重者。

1.2 方法

1.2.1 数据采集: 所有受试者均佩戴美国美敦力公司 (Medtronic, Inc) 生产的 CGMS, 该血糖仪每 10 s 接收 1 次信号, 储存每 5 min 的血糖平均值, 连续监测 72 h, 可得到 864 个血糖值。

1.2.2 建立 GM(1, 1) 模型: GM(1, 1) 模型是灰色预测模型的基本模型之一, 表达式为一阶方程、单变量。该模型在 Matlab 2018 软件环境下编码和运行, 得到未来 5、15、30 min 的血糖值。具体建模步骤^[9] 为:

实测血糖序列记为 $X^{(0)}$, 对其先进行对数变换得到 $Y^{(0)}$, 后一阶累加生成得到序列 $Y^{(1)}$, 则

$$\begin{aligned} X^{(0)} &= x^{(0)}(i), x^{(0)}(i) > 0, i = 1, 2, \dots, n \\ Y^{(0)}(i) &= \ln x^{(0)}(i), i = 1, 2, \dots, n, Y^{(0)} = y^{(0)}(n) \\ Y^{(1)}(k) &= \sum_{i=1}^k y^{(0)}(i), (k = 2, 3, \dots, n), Y^{(1)} = y^{(1)}(n) \end{aligned} \quad (1)$$

得到相应的白化微分方程, 其中 a 为发展系数, b 为灰色作用量, 表示为

$$\frac{dy^{(1)}(t)}{dt} + ay^{(1)}(t) = b, x^{(0)}(i) + az^{(1)}(i) = b, i = 2, 3, \dots, n \quad (2)$$

应用最小二乘法求出发展系数 a , 灰色作用量 b , 则

$$\hat{y}^{(0)}(i) = (1 - e^{-a}) \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(i-1)}, i = 2, 3, \dots, n \quad (3)$$

逆对数变换获得原始血糖数据的预测序列, 则

$$\begin{aligned} \hat{X}^{(0)}(i) &= e^{\hat{y}^{(0)}(i)} \\ \hat{x}^{(0)}(i) &= e^{(1-e^{-a}) \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(i-1)}}, i = 2, 3, \dots, n \end{aligned} \quad (4)$$

采用新陈代谢算法得到血糖预测值, 其原理是预测 $x^{(0)}(20)$ 时, 剔除离 $x^{(0)}(20)$ 最远的数据 $x^{(0)}(1)$, 添加离 $x^{(0)}(20)$ 最近的血糖数据 $x^{(0)}(19)$, 即以 $x^{(0)}(20) \sim x^{(0)}(19)$ 为训练集。预测 $x^{(0)}(21)$ 时, 剔除离 $x^{(0)}(21)$ 最远的数据 $x^{(0)}(2)$, 添加离 $x^{(0)}(21)$ 最近的血糖数据 $x^{(0)}(20)$, 即以 $x^{(0)}(3) \sim x^{(0)}(20)$ 为训练集。简而言之, 始终保持 18 个血糖数据的训练集, 不断更新操作。

1.2.3 GM(1, 1) 模型的预测性能评估: 通过血糖预测值与实测值的平均相对误差 (MRE)、平均绝对误差 (MAE)、均方根误差 (RMSE) 等误差分析来评价该模型的预测性能。公式如下, 其中 \hat{y}_i 代表预测血糖值, y_i 代表实测血糖值, n 为预测血糖值的数量。最后计算 50 例患者 MAE、MRE、RMSE 的平均值。

$$(1) \text{ 平均相对误差: } MRE = \frac{\sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i}}{n} \times 100\%;$$

$$(2) \text{ 平均绝对误差: } MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n};$$

$$(3) \text{ 均方根误差: } RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}.$$

1.2.4 划分不同血糖范围: 划分血糖的不同范围时, 从模型的角度出发, 将预测效率较高时的血糖范围 (目标血糖范围) 作为分组标准。应用随机数字表法抽取 20 例患者纳入对照组, 其余 30 例患者纳入实验组。对照组患者的数据用于推导分组标准, 实验组患者的数据用于统计验证。应用 Excel 2016 计算对照组所有患者预测时长为 5、15、30 min 时的 MAE。筛选血糖范围时要求连续性, 不能定义为单个或少量数据。考虑到 GM(1, 1) 模型是以 18 个连续原始血糖作为训练集进行预测, 因此要求处于目标血糖范围内的血糖数据至少连续 18 个且绝对误差均小于该预测时长的 MAE。记录符合要求的血糖范围, 求合集得到最终的目标血糖范围 (记为 $a \sim b$)。

1.2.5 分析 GM(1, 1) 模型对不同范围血糖的预测能力: 因 1 例患者连续 72 h 的血糖往往波动比较大, 很难确保均属于某一血糖范围, 故将实验

组30例患者的血糖视为1个实验单元。以对照组筛选得到的目标血糖范围将实测血糖数据分为A组(<a)、B组(a~b)、C组(>b)。预测时长为5 min时的分组记为A₁、B₁、C₁,预测时长为15 min时的分组记为A₂、B₂、C₂,预测时长为30 min时的分组记为A₃、B₃、C₃。对不同血糖组的实测值和预测值进行Pearson相关性分析评价拟合度,并应用非参数检验(多个独立样本的K-W单因素方差分析)对各组间绝对误差中位数的差异进行统计学分析。

1.3 统计学处理

本研究数据均采用SPSS 22.0软件进行统计学分析,计量资料以($\bar{x} \pm s$)表示,比较行t检验,计数资料以[n(%)]表示,比较行 χ^2 检验,P < 0.05为差异有统计学意义。

2 结果

2.1 一般资料

本研究共纳入T2DM患者50例,患者佩戴CGMS的时间至少为72 h,且血糖序列均完整连续。2组年龄、病程、性别构成、体质量指数(BMI)、糖化血红蛋白(HbA1c)比较,差异均无统计学意义(P > 0.05)。见表1。

表1 2组T2DM患者一般资料比较($\bar{x} \pm s$)[n(%)]

指标	对照组(n=20)	实验组(n=30)
年龄/岁	50.6 ± 9.4	51.0 ± 9.3
性别		
男	11(55.0)	17(56.7)
女	9(45.0)	13(43.3)
病程/年	9.7 ± 4.9	9.3 ± 4.7
BMI/(kg/m ²)	22.0 ± 1.1	22.8 ± 1.3
HbA1c/%	8.2 ± 1.2	8.5 ± 2.2

BMI: 体质量指数; HbA1c: 糖化血红蛋白。

2.2 评估模型预测性能

以患者甲为例,不同预测时长的血糖预测结果见图1。所有患者实测值与预测值的误差分析见表2,预测误差随着预测时间的延长不断增大。

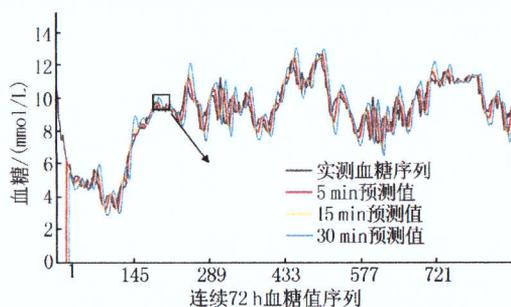


图1 患者甲不同预测时长的血糖预测结果

2.3 划分不同血糖范围

对照组预测时长为5、15、30 min时的MAE

表2 不同预测时长实测值与预测值的误差分析

误差指标	预测时长		
	5 min	15 min	30 min
MRE/%	6.10	9.00	13.10
MAE	0.46	0.68	1.01
RMSE	0.69	0.92	1.26

MRE: 平均相对误差; MAE: 平均绝对误差;

RMSE: 均方根误差。

分别为0.46、0.67、0.98。为明确绝对误差与实测血糖的变化趋势,以患者乙为例,绘制预测时长为5 min时的趋势图(见图2),然后按照绝对误差降序排列绘图(见图3)。二者对比后发现绝对误差与实测血糖变化趋势基本一致,血糖较低且平稳时预测误差小,但在打乱血糖序列的时间变化后,发现规律消失。预测时长为15、30 min时仍有相同情况,进一步证实在筛选血糖范围时要求连续性。通过筛选得到患者乙预测时长为5 min时符合要求的血糖范围有6段,求合集得到3.4~7.0、7.8~8.8、9.3~10.4 mmol/L,应用同样的方法得到预测时长为15、30 min时的目标血糖范围,见表3。对照组20例患者的结果表明,预测时长为5、15、30 min时,目标血糖范围分别为3.4~11.5、3.3~11.4、3.2~11.4 mmol/L,见表4。

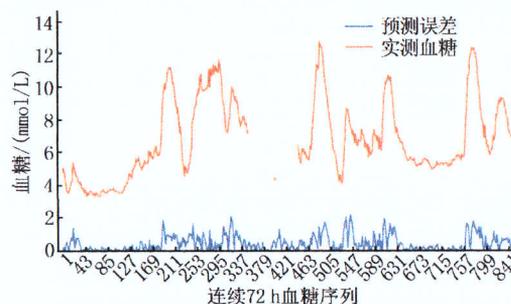


图2 患者乙预测时长为5 min时实测血糖与预测误差趋势图

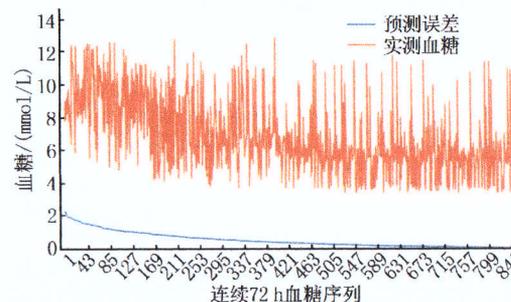


图3 患者乙预测时长为5 min时降序排列后实测血糖与预测误差趋势图

表 3 患者乙不同预测时长的目标血糖范围与最大绝对误差

项目	5 min		15 min		30 min	
	目标血糖范围/(mmol/L)	最大绝对误差	目标血糖范围/(mmol/L)	最大绝对误差	目标血糖范围/(mmol/L)	最大绝对误差
符合要求项	3.4 ~ 6.9	0.43	3.4 ~ 6.7	0.63	3.4 ~ 6.7	0.66
	9.3 ~ 10.4	0.44	7.2 ~ 11.0	0.59	9.3 ~ 10.4	0.79
	7.8 ~ 8.8	0.39	9.3 ~ 10.4	0.55	5.3 ~ 6.9	0.80
	5.3 ~ 7.0	0.36	5.3 ~ 6.9	0.45	4.8 ~ 6.2	0.61
	4.8 ~ 6.2	0.44	4.8 ~ 6.4	0.63	4.3 ~ 5.2	0.51
	4.3 ~ 5.7	0.38	4.3 ~ 5.7	0.56	5.9 ~ 8.2	0.93
合集	3.4 ~ 7.0	-	6.4 ~ 8.2	0.45	3.4 ~ 8.2	-
	7.8 ~ 8.8				9.3 ~ 10.4	
	9.3 ~ 10.4					

表 4 对照组患者不同预测时长的目标血糖范围与最大绝对误差

患者编号	目标血糖范围/(mmol/L)			最大绝对误差		
	5 min	15 min	30 min	5 min	15 min	30 min
1	3.4 ~ 6.5	3.4 ~ 7.4	3.4 ~ 7.4	0.41	0.61	0.86
2	4.9 ~ 7.1	4.9 ~ 7.1	4.9 ~ 7.3	0.42	0.59	0.94
3	5.5 ~ 8.7	5.5 ~ 8.7	5.8 ~ 8.7	0.41	0.61	0.94
4	3.4 ~ 4.6	3.4 ~ 4.6	3.2 ~ 4.9	0.44	0.64	0.97
	5.4 ~ 8.7	5.4 ~ 8.7	5.4 ~ 8.7			
5	5.4 ~ 7.8	5.4 ~ 7.7	5.4 ~ 7.7	0.45	0.65	0.75
6	6.7 ~ 7.5	6.3 ~ 7.5	6.3 ~ 7.5	0.41	0.46	0.66
7	3.4 ~ 7.0	3.4 ~ 11.0	3.4 ~ 6.9	0.44	0.63	0.80
	7.8 ~ 8.8		9.3 ~ 10.4			
	9.3 ~ 10.4					
8	5.6 ~ 8.0	3.3 ~ 4.9	3.3 ~ 5.0	0.42	0.61	0.93
		5.6 ~ 7.8	5.5 ~ 7.8			
9	4.2 ~ 7.4	4.3 ~ 7.4	4.2 ~ 7.9	0.39	0.63	0.95
10	4.5 ~ 5.8	4.5 ~ 6.0	4.5 ~ 6.0	0.42	0.57	0.83
	7.8 ~ 10.2	7.9 ~ 10.2	8.2 ~ 9.7			
	10.9 ~ 11.4	10.9 ~ 11.4	10.9 ~ 11.4			
11	8.2 ~ 9.2	8.2 ~ 11.3	8.2 ~ 10.8	0.44	0.65	0.92
12	5.2 ~ 6.9	5.2 ~ 6.9	5.2 ~ 6.9	0.43	0.53	0.68
13	6.1 ~ 9.0	6.1 ~ 8.5	5.6 ~ 8.2	0.42	0.65	0.91
14	4.7 ~ 7.7	4.7 ~ 7.0	4.6 ~ 7.5	0.39	0.60	0.89
15	4.4 ~ 8.2	4.4 ~ 10.0	4.4 ~ 9.9	0.42	0.61	0.90
	8.7 ~ 10.0					
16	3.4 ~ 6.2	3.4 ~ 6.2	3.4 ~ 7.9	0.43	0.57	0.89
	6.9 ~ 8.7	6.9 ~ 8.6				
17	5.3 ~ 8.5	5.3 ~ 8.5	5.2 ~ 8.7	0.44	0.62	0.94
18	6.7 ~ 9.9	6.3 ~ 11.3	6.2 ~ 11.3	0.44	0.66	0.92
	10.5 ~ 11.5					
19	6.1 ~ 7.0	6.1 ~ 7.0	6.5 ~ 7.7	0.45	0.63	0.94
	8.1 ~ 10.6	8.1 ~ 10.6	8.1 ~ 10.6			
20	5.1 ~ 8.4	5.1 ~ 8.5	5.1 ~ 8.3	0.45	0.62	0.83
合集	3.4 ~ 11.5	3.3 ~ 11.4	3.2 ~ 11.4	-	-	-

2.4 分析 GM(1, 1) 模型在不同血糖范围的预测能力
 预测时长为 5、15、30 min 时分别得到 25 380、

25 320 及 25 230 对实测血糖值与预测血糖值。不同预测时长的绝对误差近似正态分布 ($P > 0.05$), 相关性分析表明, 不同预测时长时, B 组

的 r 值均最大 ($P < 0.01$), C 组 r 值次之, A 组 r 值最小, 见表 5。为进一步明确各组间 MAE 的差异, 进行单因素方差分析, 发现方差不齐 ($P < 0.05$), 故采用非参数检验进行分析。结果表明, 每 2 组之间的绝对误差中位数比较, 差异有统计学意义 ($P < 0.01$), 其中 A 组的预测误差最小, B 组次之, C 组最大, 见表 6。因此, 预测时长为 5、15、30 min 时, GM(1, 1) 模型分别对处于 3.4 ~ 11.5、3.3 ~ 11.4、3.2 ~ 11.4 mmol/L 血糖的预测拟合度最好且预测误差较小。

表 5 不同预测时长、不同血糖范围的相关性分析及 MAE 结果

预测时长	组别	对数	r	P	MAE
5 min	A ₁ 组	324	0.468	<0.001	0.37
	B ₁ 组	23 082	0.937	<0.001	0.44
	C ₁ 组	1 974	0.873	<0.001	0.87
15 min	A ₂ 组	290	0.375	<0.001	0.51
	B ₂ 组	22 967	0.875	<0.001	0.64
	C ₂ 组	2 063	0.770	<0.001	1.34
30 min	A ₃ 组	255	0.189	0.002	0.71
	B ₃ 组	22 925	0.773	<0.001	0.94
	C ₃ 组	2 050	0.631	<0.001	2.07

A₁ 组: 血糖 <3.4 mmol/L 组; B₁ 组: 血糖 3.4 ~ 11.5 mmol/L 组;
C₁ 组: 血糖 >11.5 mmol/L 组; A₂ 组: 血糖 <3.3 mmol/L 组;
B₂ 组: 血糖 3.3 ~ 11.4 mmol/L 组; C₂ 组: 血糖 >11.4 mmol/L 组;
A₃ 组: 血糖 <3.2 mmol/L 组; B₃ 组: 血糖 3.2 ~ 11.4 mmol/L 组;
C₃ 组: 血糖 >11.4 mmol/L 组; MAE: 平均绝对误差。

表 6 不同预测时长、不同血糖范围的非参数检验

预测时长	组别	绝对误差 中位数	多个独立样本的非参数检验	
			组间比较	P
5 min	A ₁ 组	0.21	A ₁ 组与 B ₁ 组	0.001
	B ₁ 组	0.27	B ₁ 组与 C ₁ 组	<0.001
	C ₁ 组	0.61	A ₁ 组与 C ₁ 组	<0.001
15 min	A ₂ 组	0.30	A ₂ 组与 B ₂ 组	<0.001
	B ₂ 组	0.39	B ₂ 组与 C ₂ 组	<0.001
	C ₂ 组	0.93	A ₂ 组与 C ₂ 组	<0.001
30 min	A ₃ 组	0.39	A ₃ 组与 B ₃ 组	<0.001
	B ₃ 组	0.57	B ₃ 组与 C ₃ 组	<0.001
	C ₃ 组	1.38	A ₃ 组与 C ₃ 组	<0.001

A₁ 组: 血糖 <3.4 mmol/L 组; B₁ 组: 血糖 3.4 ~ 11.5 mmol/L 组;
C₁ 组: 血糖 >11.5 mmol/L 组; A₂ 组: 血糖 <3.3 mmol/L 组;
B₂ 组: 血糖 3.3 ~ 11.4 mmol/L 组; C₂ 组: 血糖 >11.4 mmol/L 组;
A₃ 组: 血糖 <3.2 mmol/L 组; B₃ 组: 血糖 3.2 ~ 11.4 mmol/L 组;
C₃ 组: 血糖 >11.4 mmol/L 组。

3 讨论

为实现糖尿病患者智能化控制血糖以及更好地改善预后, 人工胰腺 (AP) 概念被提出, 现 AP 已被允许使用于临床^[10-12]。多项研究表明 AP

或闭环胰岛素泵应用于不同类型糖尿病中均相对安全, 其中一项研究^[13]证实了 AP 治疗 1 型糖尿病 (T1DM) 患者的有效性和安全性。1 例首次怀孕时使用胰岛素皮下治疗而第 2 次怀孕时使用 AP 治疗的糖尿病患者案例^[14]发现, AP 可使血糖管理更容易。另一项针对住院 T2DM 患者的研究^[15]发现, 闭环胰岛素泵输注组的目标范围血糖比例明显高于常规胰岛素皮下输注组。虽然 AP 目前仍存在伦理^[16]、传感器性能^[17]等方面问题, 但其对糖尿病患者的益处可能更多。在构建闭环装置时, 血糖的控制算法是不可缺少的一环^[18], 但也带来了重大挑战, 尤其是在不规律的食物摄入、运动和各种自发活动下使用机器学习算法实现血糖预测。

在 CGMS 越来越成熟的基础上, 学者们致力于探索不同模型在血糖预测中的应用效果。一种是基于生理学的模型^[19] (葡萄糖-胰岛素代谢模型和葡萄糖吸收模型), 由于需要考虑详细的生理过程, 建模复杂耗时, 存在操作困难的问题。另一种是数据驱动模型^[20], 其基于血糖数据的数学定律建立模型, 操作更容易, 且预测精度良好。灰色预测模型通过模糊数学处理灰色信息逐渐发现系统中的未知信息, 具有较少的训练集即可实现高精度的优势^[8], 且建模步骤简单易操作, 这也是本研究选择该模型的原因。灰色预测模型应用于血糖预测中对糖尿病患者具有重要价值, 当预测得知患者即将发生低血糖时, 可提示患者补充糖分; 当预测得到的血糖值偏高时, 可指导临床医生调整胰岛素用量以维持患者血糖稳定。

GM(1, 1) 模型是灰色预测理论中应用最广泛的动态预测模型之一, 其本质属于指数预测模型, 因此其预测精度与数据序列的平滑程度以及被预测对象的跃迁变化规律密切相关。GM(1, 1) 模型的核心是通过灰色微分拟合方法建立离散拟合方程, 但该方程是近似微分方程, 因此很难保证模型的固有误差必然是无穷小的。为了减少由建模方法缺陷而导致的固有误差, 本研究对传统的 GM(1, 1) 模型进行了改进。首先是数据的预处理, 即对实测血糖序列取对数, 这样可以提高数据的平滑度。其次, 利用最小二乘法求解 GM(1, 1) 模型的参数, 从而提高拟合方程与待拟合方程之间的近似性。再者, 为了延长预测时间, 预测的序列被横向延长 1 步、3 步和 6 步获得不同的预测时间。最终, GM(1, 1) 模型通过揭示血糖中蕴

含的数学规律,拟合下一时刻的血糖值。根据实验结果,GM(1, 1)模型应用于血糖预测时,预测时间越短,预测性能越好,预测时间延至 30 min 时,平均绝对误差达到 1.01,仍可接受,这也证实了 GM(1, 1)模型对血糖的短期预测是有效且准确的。

既往血糖预测模型的应用研究多局限于比较不同模型的精确度^[6-9]、某一模型的性能评估^[4-5, 10],并未对某一确定模型对不同范围血糖的预测能力行进一步讨论。另外团队前期工作发现,GM(1, 1)模型对处于 3.9 ~ 10.0 mmol/L 的血糖预测拟合度较好($r=0.85$, $P<0.01$),但却是根据动态血糖监测仪设定的血糖水平直接进行分组分析,并未从模型角度去细化评价 GM(1, 1)模型对哪一范围血糖具有最佳的预测效果。因此,本研究根据筛选得到的血糖范围进行分组,增强了血糖范围与模型的适配度,旨在为将模型更精确地应用于临床提供一定参考。

Pearson 相关性分析及非参数检验结果表明,GM(1, 1)模型对不同范围血糖的预测能力是有差异的。以预测时长 5 min 的结果为例,从预测拟合度来看,GM(1, 1)模型对处于 3.4 ~ 11.5 mmol/L 范围血糖的预测效果最好,其次是 > 11.5 mmol/L 的血糖,对 < 3.4 mmol/L 的血糖预测效果最差。灰色预测模型对时间序列(如血糖序列)进行预测时,抛去模型的固有误差,数据的随机波动是影响预测的主要因素^[11],即当被预测对象越平滑、随机性越弱,其预测效果越好。临床患者真实发生的低血糖事件较少,且低血糖持续时间短,同样高血糖的发生常伴随着机体自身调节或临床治疗措施所致的短期内血糖下降,这也造成高血糖的时间比例相对较少,而在高血糖与低血糖发生前后的数据均被纳入目标血糖组。因此,从数据的连续性、平滑度来说,处于目标血糖范围的数据的连续性最好且随机变化的值较少,高血糖组次之,低血糖组最差,这就合理解释了 GM(1, 1)模型在目标血糖范围表现出的最佳预测拟合性。从预测误差来看,预测时长为 5 min 时,GM(1, 1)模型对 < 3.4 mmol/L 的血糖预测误差最小,处于 3.4 ~ 11.5 mmol/L 范围的血糖次之, > 11.5 mmol/L 的血糖误差最大。MAE 主要反映预测值与实际值之间的差距,取决于二者差值的绝对值大小。高血糖发生后往往有血糖的下降,由于 GM(1, 1)模型是基于前 18 个训练集的数据进行预测的,包

含高血糖极值的训练集得到的预测值往往会偏高,而所对应的真实值已存在一定程度下降,二者之间的绝对误差较大,这可能是高血糖组预测误差最大的主要原因。而低血糖组由于可变化范围窄、波动幅度小,预测值与实测值之间的差值最小。总之,本研究通过对不同预测时长的血糖范围求合集,得出 GM(1, 1)模型对处于 3.4 ~ 11.4 mmol/L 的血糖数据预测更有效。

如果某预测模型对不同范围的血糖都表现出良好的预测性能,则其更具有临床适用性。在众多血糖预测模型中合理选择模型以实现精准预测血糖,对于延缓糖尿病并发症的发生与发展有着重要意义。GM(1, 1)模型可通过对少量、杂乱、不确定信息的数据建立灰色微分方程,描述事物进一步的发展规律。本研究结果显示,GM(1, 1)模型能有效预测血糖,其预测特点是对 3.4 ~ 11.4 mmol/L 范围的血糖数据预测效果最好,这不仅提示了该模型进行血糖预测时的建议适用范围,也提示在对发生高血糖较多的波动性血糖序列进行预测时,将该模型与其他模型组合预测可能会使预测值更准确。然而,GM(1, 1)模型仍然存在对偏高血糖、极低血糖预测效果不佳的局限,这也是临床研究人员未来面临的新挑战。

参考文献

- [1] AKBAS E M, DEMIRTAS L, OZCICEK A, *et al.* Association of epicardial adipose tissue, neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio with diabetic nephropathy[J]. *Int J Clin Exp Med*, 2014, 7(7): 1794 - 1801.
- [2] 中华医学会糖尿病学分会. 中国 2 型糖尿病防治指南(2017 年版)[J]. *中国实用内科杂志*, 2018, 38(4): 292 - 344.
- [3] Seuring T, Archangelidi O, Suhreke M. The economic costs of type 2 diabetes: a global systematic review[J]. *Pharmacoeconomics*, 2015, 33(8): 811 - 831.
- [4] SATTI K, MACAŠ M, LHOTSKÁ L, *et al.* Ensemble methods in combination with compartment models for blood glucose level prediction in type 1 diabetes mellitus[J]. *Comput Methods Programs Biomed*, 2020, 196: 105628.
- [5] HE J, WANG Y. Blood glucose concentration prediction based on kernel canonical correlation analysis with particle swarm optimization and error compensation[J]. *Comput Methods Programs Biomed*, 2020, 196: 105574.
- [6] 都承华, 龚谊承. 基于支持向量回归和核岭回归对血糖值预测的对比分析[J]. *数学的实践与认识*, 2020, 50(6): 132 - 139.
- [7] 王延年, 郭占丽, 袁进磊, 等. 基于自适应噪声完整聚类经验模态分解-极限学习机的短期血糖预测[J]. *中国生物医学工程学报*, 2017, 36(6): 702 - 710.

- [4] 涂业桃. 抗-M 抗体对 ABO 血型鉴定的影响[J]. 国际检验医学杂志, 2016, 37(17): 2461-2463.
- [5] 陈倩, 唐秋萍. 105186 份无偿献血者不规则抗体筛查结果分析[J]. 热带医学杂志, 2020, 20(9): 1248-1250, 1258.
- [6] 郑楚忠, 唐万兵, 郑望春, 等. 深圳地区无偿献血者不规则抗体筛查结果分析[J]. 国际检验医学杂志, 2015, (12): 1693-1694, 1697.
- [7] 靖春旭, 沈延平. 南阳地区无偿献血者血液标本红细胞血型不规则抗体筛查结果分析[J]. 河南预防医学杂志, 2020, 31(2): 134-136.
- [8] 董珀, 濮云峰, 许飞. 2015—2018 年常州地区献血者不规则抗体的检出情况分析[J]. 临床血液学杂志, 2019, 32(10): 793-795.
- [9] 丁梦圆, 潘志荣, 金一鸣, 等. 苏州市无偿献血者与 O 型反定型细胞反应的抗体筛查分析[J]. 临床血液学杂志: 输血与检验, 2019, 32(5): 740-743.
- [10] 蔡茵, 曹斌, 周喆, 等. 上海地区无偿献血者血液中不规则抗体的调查[J]. 中国输血杂志, 2017, 30(7): 791-793.
- [11] 詹廷西, 王静, 徐丹丹, 等. 输血前患者不规则抗体筛查及鉴定的临床意义探讨[J]. 中国免疫学杂志, 2021, 37(3): 335-338, 345.
- [12] 付蒙, 张京, 曹冬如, 等. 孕产妇不规则抗体筛查的临床价值和意义[J]. 中国临床医生杂志, 2018, 46(1): 96-98.
- [13] 石祖亮, 杨林, 刘贲, 等. 武汉地区 54524 例孕产妇不规则抗体筛查结果分析[J]. 临床血液学杂志, 2017, 30(6): 449-451.
- [14] 戎霞, 张雪梅, 黄珂, 等. 广州地区无偿献血者不规则抗体筛查结果分析[J]. 热带医学杂志, 2014, 14(2): 176-178, 183.
- [15] 郑妍, 王文婷, 穆士杰. 血型不规则抗体阳性患者的管理[J]. 国际检验医学杂志, 2018, 39(1): 125-127.

(本文编辑: 陆文娟)

(上接第 28 面)

- [8] 邓聚龙. 灰色系统理论教程[M]. 武汉: 华中理工大学出版社, 1990: 122-123.
- [9] 覃伟, 高敏, 沈莹, 等. 基于机器学习算法的 2 型糖尿病患者 3 个月血糖预测[J]. 中华疾病控制杂志, 2019, 23(11): 1313-1317.
- [10] WANG Y, WEI F, SUN C, *et al.* The Research of Improved Grey GM(1, 1) Model to Predict the Postprandial Glucose in Type 2 Diabetes [J]. *Biomed Res Int*, 2016, 2016: 6837052.
- [11] 龚谊承, 都承华, 张艳娜, 等. 基于主成分和GBDT对血糖值的预测[J]. 数学的实践与认识, 2019, 49(14): 116-122.
- [12] VOELKER R. "artificial pancreas" is approved[J]. *JAMA*, 2016, 316(19): 1957.
- [13] KOVATCHEV B, CHENG P, ANDERSON S M, *et al.* Feasibility of long-term closed-loop control: a multicenter 6-month trial of 24/7 automated insulin delivery[J]. *Diabetes Technol Ther*, 2017, 19(1): 18-24.
- [14] SCHÜTZ-FUHRMANN I, SCHÜTZ A K, EICHNER M, *et al.* Two subsequent pregnancies in a woman with type 1 diabetes: artificial pancreas was a gamechanger[J]. *J Diabetes Sci Technol*, 2020, 14(5): 972-973.
- [15] THABIT H, HARTNELL S, ALLEN J M, *et al.* Closed-loop insulin delivery in inpatients with type 2 diabetes: a randomised, parallel-group trial [J]. *Lancet Diabetes Endocrinol*, 2017, 5(2): 117-124.
- [16] QUINTAL A, MESSIER V, RABASA-LHORET R, *et al.* A critical review and analysis of ethical issues associated with the artificial pancreas [J]. *Diabetes Metab*, 2019, 45(1): 1-10.
- [17] RAMLI R, REDDY M, OLIVER N. Artificial pancreas: current progress and future outlook in the treatment of type 1 diabetes [J]. *Drugs*, 2019, 79(10): 1089-1101.
- [18] BLAUW H, KEITH-HYNES P, KOOPS R, *et al.* A review of safety and design requirements of the artificial pancreas [J]. *Ann Biomed Eng*, 2016, 44(11): 3158-3172.
- [19] QIN Y F, XIAO J, CHEN X D. Physiological model and chemical modeling strategy for blood glucose prediction [J]. *Chemical Industry and Engineering Progress*, 2019, 38(1): 545-555.
- [20] OVIEDO S, VEHI J, CALM R, *et al.* A review of personalized blood glucose prediction strategies for T1DM patients [J]. *Int J Numer Meth Biomed Engng*, 2017, 33(6): e2833.

(本文编辑: 陆文娟)