International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:3, No:9, 2009

Agent-based Simulation for Blood Glucose Control in Diabetic Patients

Sh. Yasini, M. B. Naghibi-Sistani, A. Karimpour

***Abstract*—**This paper employs a new approach to regulate the blood glucose level of type I diabetic patient under an intensive insulin treatment. The closed-loop control scheme incorporates expert knowledge about treatment by using reinforcement learning theory to maintain the normoglycemic average of 80 mg/dl and the normal condition for free plasma insulin concentration in severe initial state. The insulin delivery rate is obtained off-line by using Q-learning algorithm, without requiring an explicit model of the environment dynamics. The implementation of the insulin delivery rate, therefore, requires simple function evaluation and minimal online computations. Controller performance is assessed in terms of its ability to reject the effect of meal disturbance and to overcome the variability in the glucose-insulin dynamics from patient to patient. Computer simulations are used to evaluate the effectiveness of the proposed technique and to show its superiority in controlling hyperglycemia over other existing algorithms.

***Keywords*—**Insulin Delivery rate, Q-learning algorithm, Reinforcement learning, Type I diabetes.

I. Introduction

HUMAN bodies need to maintain glucose concentration level in a narrow range 70-110 mg/dl. If one's glucose concentration level is significantly out of the normal range, this person is considered to have the plasma glucose problem: Hyperglycemia or hypoglycaemia.

Diabetes mellitus is a disease in glucose-insulin endocrine metabolic system, in which the pancreas either does not release insulin or does not properly use insulin to uptake glucose in the plasma, which is referred as hyperglycemia [1]. The two types of diabetes are Type I and Type II. In this paper the focus is on type I diabetes. In Type I diabetes, the body's immune system destroys pancreatic beta cells, and the patient is totally dependent on an external source of insulin to be infused at an appropriate rate to maintain the blood glucose concentration.

When a normal person is subjected to a glucose meal, the glucose concentration in plasma increases from basal value and so the pancreatic β-cells secrete insulin. The insulin in plasma is hereby increased, and the glucose uptake in muscles, liver, and tissues is raised by the remote insulin in action. This lowers the glucose concentration in plasma, implying the β-cells to secrete less insulin, from which a feedback effect arises [2]. But, in type I diabetic patients whose pancreas does not release insulin, blood glucose level remains in much more than basal value for long period of time. When glucose level remains high for extended periods of time the patient is at risk for neuropathy, nephropathy, blindness, and other long-term vascular complications. However, the result of the Diabetes Control and Complications Trial (DCCT) showed that an intensive insulin therapy can reduce the risk of developing complications [3]. Consequently, an intensive therapy is encouraged for type I diabetic patients prescribed by a continuous subcutaneous insulin infusion pump.

Control strategies of diabetes treatment can be categorized as open loop control, semi closed-loop, and closed-loop control. Current treatment methods utilizing open loop control in which physicians inject a pre determined dose of insulin subcutaneously based on three or four time daily glucose measurements, usually by an invasive method of finger prick. This method not only is painful and inconvenient but also unreliable because of approximation involved in type and the amount of insulin delivered. In semi closed-loop control insulin infusion rate adjust according to intermittent blood glucose readings. This technique is sub-optimal and unable to accomplish the aforementioned normalization and also suffered from long sampling time problem of missing fast or inter-sample disturbances. However, closed-loop control method which acts as an artificial pancreas is the most effective way of diabetes treatment and could improve the quality of life and life expectancy of patients [4]. Ultimately, a true artificial pancreas is a closed-loop device that enables a person with diabetes to maintain normal glucose levels by providing the right amount of insulin at the right time, just as the pancreas does in non-diabetic individuals [5].

In the near term, we expect artificial pancreases to be external devices comprises of insulin pumps, already widely available; continuous glucose monitors (CGMs), which are coming on the market now and an appropriate control algorithm. Figure 1 shows the block diagram of a closed-loop control system of diabetic patients. In this system, the control algorithm would calculate optimal insulin delivery rate designed to keep the patient under metabolic control, and a signal would drive a mechanical pump to deliver the desired amount of insulin.

Sh. Yasini, is with the Electrical engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran (corresponding author to provide phone: 98-9155404475; e-mail: sh_ya85@stu-mail.um.ac.ir).

M. B. Naghibi-Sistani is with department of electrical engineering, Ferdowsi University of Mashhad, Mashhad, Iran (e-mail: naghib@yahoo.com).

A. karimpour is with the Electrical Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran (e-mail: karimpor@um.ac.ir).

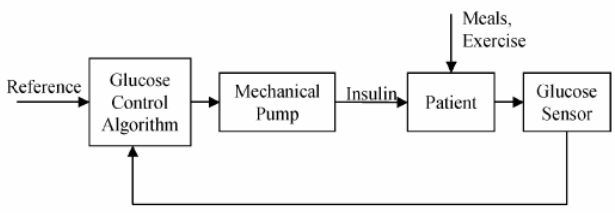

Fig. 1 Closed-loop control of diabetic patient

Since recent advances have made available programmable and variable-rate infusion pumps, the feedback control system mimics the normal function of a pancreas more closely. However, creating a device which would accurately replace multiple insulin injections per day for a long period of life is not an easy task. It should be made from biocompatible materials and as small as possible. Four major sites for invasive insulin delivery are subcutaneous, intramascular, intravenous, and intrperitoneal [6]. The subcutaneous site is the simplest and safest in long term but the absorption of insulin from subcutaneous tissue is delayed. The intramascular site is usually preferred for people affected by brittle diabetes who have a subcutaneous barrier to insulin absorption. The intravenous has rapid delivery with negligible dead-time. The main problem of this approach is presence of the intravenous lines which may not be suitable for some patients. Intraperitoneal is the most physiological insulin delivery, though the major disadvantage is its difficult access. The recent advances have brought in non-ivasive modes of insulin delivery such as transdermal and oral [7]-[8]. These modes are not painful like the invasive modes but have problems such as low skin permeability in transdermal mode and issues concerned with the oral bioavailability for the oral mode.

Continuous glucose monitors are devices that provide continuous "real time" readings and data about trends in glucose levels. Blood glucose monitoring devices are classified as invasive, minimally invasive, and non-invasive. Fully invasive systems can be either beside clinical devices or self-monitoring meters. Such system allows continuous monitoring, therefore increasing the amount of clinical information. System which puncture the skin are still standard techniques for home monitoring reading glucose concentration through electrochemical or optical disposable strips for finger prik blood samples [9]. Efforts have been made to reduce the level of invasiveness by decreasing the blood sample volume to a few microliters, and measuring areas of the body less sensitive to pain than fingertrips such as forearm, upper arm, or thigh. Minimally invasive measurements sample the interstitial fluid with subcutaneous sensors [10]. Even in this method the discomfort causes difficulties to the patient's therapy. Therefore, researches group are working to develop non-invasive glucose control devices [11].

In testing the performance of the control algorithm a virtual patient need to be implemented using an appropriate mathematical model. During the last decades, many mathematical models have been derived to describe dynamics of glucose-insulin regulatory system [12]-[14]. These models have ranged from linear to nonlinear with increasing the levels of complexity [15]. Since, the parameters of these models are in general time-varying, even for a patient under a constant treatment and environment conditions. Therefore, employed controller in closed-loop system should be robust to parameters variations in model and physical disturbances like food intake.

With the availability of these mathematical models different algorithms based on control theory have been developed to control the blood glucose level in people with diabetes. Some of these algorithms include proportional-integral-derivative (PID) [16], [17] and proportional-derivative (PD) [18], that need a linearized model for the design, as well as H_∞ control technique. If linear models are employed for the patients, control algorithm like H_∞ control can guarantee some level of performance but full robustness can not be achieved via this algorithms. However, as far as linear control algorithms are concerned, H_∞ control offers a promising result in maintaining blood glucose regulation in diabetic patients. Some interesting result of this method can be found in [19]-[20]. Also optimal control algorithms are applied for blood glucose regulation in semi closed-loop control system [22], [23]. But, the important point in most of these researches is that proposed controller has been designed with regard to mathematical model as a crisp model, and uncertainty in the model parameters has been not considered. Therefore, although these methods, would offer good responses in simulations, it is likely that they would not be successful in practice and failed while applying to an actual patient. On the other hand, biomedical systems are inherently complex and nonlinear, and are often correlated with imprecision and model parameter variations. Consequently, conventional control techniques can prove insufficient for controlling such systems.

The ultimate goal of this research is to develop a consistent, robust controller for safe, predictable regulation of blood glucose levels in diabetic patients.

In control theory, reinforcement learning has emerged as a powerful tool to incorporate knowledge about the system for implementing an appropriate closed-loop control law without requiring an explicit model of the environment. This approach is based on trial and error search [24]. Thus, the agent can learn from its mistakes and adapt his intelligent treatments to uncertainty and variations in the environment.

This work exploits agent-based simulation under Q-learning algorithm to regulate the blood glucose level of type I diabetic patient around euglycemia. Insensitive to disturbance, accuracy, and robustness to uncertainty as well as appropriate settling time are main features of proposed algorithm. The text is organized as follows. The physiological model of glucose-insulin regulatory system in type I diabetes mellitus patient is introduced in section 2. In section 3 some fundamentals on reinforcement learning are presented along with a description of the Q-learning algorithm. Simulation results and concluding remarks are included in sections 4 and 5, respectively.

II. Glucose Insulin Regulatory System Dynamic Model

Complex models though are accurate for regimen evaluation but are generally unsuited for real-time control due to they need several time points of input to produce the insulin infusion profile. Additionally, they are not generic requiring the data of a specific patient and known glucose inputs. Against, simple models capture essential dynamics behaviors and provide a more suitable foundation for real-time control design.

The aim of this paper is to develop a control technique based on a physiological model that capture the essential system dynamics, which do not require unavailable data, and are applicable to a wider variety of subjects. Simple models capture these essential dynamic behaviors, providing a more suitable model for real-time control design and analysis.

Bergman's minimal model has proposed as a powerful modeling approach to estimating the insulin sensitivity and the glucose effectiveness, which are very useful in the study of diabetes and is the most popularly used model in the literature which has the following advantages [25], [26]:

- to be physiologically based,
- having parameters that can be estimated with a reasonable precision,
- parameters with values that are reasonable and have physiological interpretation,
- best able to simulate the dynamics of the system with the smallest number of identifiable parameters

The third-order model is comprised of a glucose compartment, G , a remote insulin compartment, X , and an insulin compartment, I . The remote insulin compartment mediates glucose uptake within the glucose space to the peripheral and hepatic tissues. The model equations are [25]:

$$\begin{aligned} \dot{G}(t) &= -p_1[G(t)-G_b]-X(t)G(t)+D(t), \quad G(0)=G_0 \\ \dot{X}(t) &= -p_2X(t)+p_3[I(t)-I_b], \quad X(0)=0 \\ \dot{I}(t) &= -n[I(t)-I_b]+\gamma[G(t)-h]^+t+u(t), \quad I(0)=I_0 \end{aligned} \tag{1}$$

Where t=0 is the glucose injection, + denotes positive reflection and:

$G(t)$: the plasma glucose concentration at time t (mg/dl), $X(t)$: is the generalized insulin variable for the remote compartment (min^{-1}), $I(t)$: is the plasma insulin concentration at time t (μU/ml), G_b: is the basal preinjection value of plasma glucose (mg/dl), I_b: is the basal preinjection value of plasma insulin (μU/ml). p_1: insulin independent rate constant of glucose rate uptake in muscles, liver and adipose tissue (min^{-1}), p_2: the rate of decrease in tissue glucose uptake ability (min^{-1}), p_3: the insulin independent increase in glucose uptake ability in tissue per unit of insulin concentration above I_b (min^{-2}(μU/ml)), n: the first order decay rate for insulin in plasma (min^{-1}), h: the threshold value of glucose above which the pancreatic β-cells release insulin, γ: the rate of the pancreatic β-cells' release of insulin after the glucose injection and with glucose concentration above h [(μU/ml) min^{-2} $(mg/dl)^{-1}$], G_0: the theoretical glucose concentration in plasma (mg/dl) at time 0, I_0: the theoretical insulin concentration in plasma (μU/ml) at time 0.

The term $\gamma[G(t)-h]^+$ in the third equation of the model acts as an internal regulatory function that formulates the insulin secretion in the body, which does not exist in diabetic patients. The metabolic portrait of a single individual is then determined by the following parameters:

$$\text{Insulin Sensitivity: } S_I = \frac{p_3}{p_2} \tag{2}$$

$$\text{Glucose Effectiveness: } S_G = p_1 \tag{3}$$

$$\text{Pancreatic responsiveness: } \varphi_1 = \frac{I_{\max}-I_b}{n(G_0-G_b)}, \quad \varphi_2 = \gamma \times 10^4 \tag{4}$$

Where I_{max} is the maximum value of insulin in plasma. S_I is measured in $(\mu U/ml)^{-1}$ per minute. S_G in min^{-1} and φ_1 in $min^{-1}\mu U/ml$ per mg/dl. These factors are important indicatives of how glucose and insulin act inside that person's body.

The available clinical data indicates that the value of p_1 parameter for diabetic patient will be significantly reduced and it can be approximated as zero [23]. Model parameters and constants are adopted from [23], [26] are given in table 1. Note that these values were calculated for a person of average weight and vary from patient to patient which makes the design of controller a more challenging task.

TABLE I
MODEL PARAMETERS

Parameter	Value
p_1	0.0316
p_2	0.0107
p_3	5.3×10^{-6}
n	0.2640
h	80.2576
γ	0.0042
G_b	70
I_b	7

$D(t)$ shows the meal glucose disturbance and can be modeled by decaying exponential function of the following form [23]:

$$D(t) = A\exp(-Bt), \quad B > 0 \tag{5}$$

Where t is in min and $D(t)$ is in (mg/dl/min). $u(t)$ is the exogenous insulin infusion rate. The model is simple, yet accurately represents the essential dynamics of the human

glucose-insulin regulatory system. The controller uses a feedback loop that employs the blood glucose level G, and its derivative (dG/dt), as sensor inputs, and the exogenous insulin infusion rate $u(t)$ as the control output.

III. Introduction to Reinforcement Learning Theory and Q-Learning Algorithm

A. Reinforcement learning

Many learning theories have been developed as a result of man's effort to analyze the behavior of animals and artificial systems. Reinforcement Learning (RL) is one of them and focuses on the effect of rewards and punishments on subjects choices in their attempt to achieve a goal; it studies complex behaviors, where sometimes taking an unpleasant action may lead to a long term reward [24]. The basic elements of RL, are;

- The learner or the decision maker, called the *agent*,
- Everything it interacts with, called the *environment*.

The learning process is described in figure 2. At each time step, t. the agent receives the state, x_t, of its environment and selects an action, a_t, according to this perception and to its past experience. One time step later, in part as a consequence of its action, the agent receives a numerical reward, r_t, and finds itself in a new state. At each time step, the agent implement a mapping from state representation to probabilities of selection each possible action. The mapping is called the agent's *policy*.

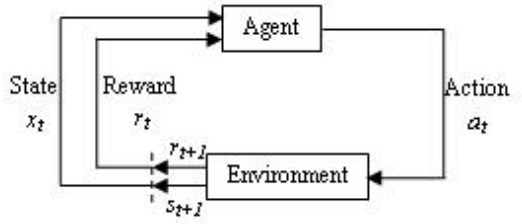

Fig. 2 The learning process of the agent through its interaction with the environment

As it is obvious from the process described before, the two basic concepts behind RL are trial and error search, since the agent explores its environment and learns from its mistakes.

B. Q-learning Algorithm

Significant advance in the field of reinforcement learning is the Q-learning algorithm of Watkins, 1998 [27]. Its main advantages are that it can be used online without having an explicit model of the environment.

As shown in figure 2, a numerical reward r_t corresponds to each pair (x_t,a_t). Therefore the reward is a function of the state received by the agent and the action it takes. The agent in Q-learning keeps in memory a function $Q_t(x_t,a_t)$ that represent the expected payoff it believes it will obtain by taking an action. The function of the expected reward is represented by a two-dimensional lookup table indexed by state-action pairs, whose elements are defined as Q-values.

The agent experience, concerning its interaction with the environment, consists of a sequence of distinct stages. Let $X=\{x_1,x_2,\ldots,x_k\}$ be the set of k possible states of the environment and $A=\{a_1,a_2,\ldots a_m\}$ be the set of m possible actions the agent can take. In the nth episode, the agent:

1. Observes its current state $x_t \in X$.
2. Selects and performs an action $a_t \in A$ using a policy.
3. Observes the subsequent state $x_{t+1} \in X$.
4. Receives a numerical reward r_t.
5. Update its Q values according to:

$$Q_t(x,a) = (1-\alpha_t)Q_{t-1}(x,a) + \alpha_t[r_t + \gamma V_{t-1}(x_{t+1})] \quad (6)$$

$$V_{t-1}(x+1) \equiv \max_b \{Q_{t-1}(x+1,b)\} \quad (7)$$

Equation (6) denotes the best the agent thinks it can do in state x_t. According to equation (7), only Q values corresponding to current state and last action chosen are updated. $\pi_n(x,a)$ is a learning rate in the range (0,1], that reflects the degree to which estimated Q values are updated by new data. α is learning rate in the range [0,1), that reflects the degree to which estimated Q values are updated by new data. γ is a discount factor in the range [0,1), representing the weight given to future reinforcements, the closer γ is to 1 the more important are distant payoffs, and, typically, the more difficult the optimization problem [24], [27].

C. Q-learning Implementation for blood Glucose regulation

The purpose of applying Q-learning algorithm for diabetes control is to determine appropriate insulin infusion rate in order to stabilize blood glucose level of diabetic patient in a reasonable time interval. Q-learning algorithm application in developing diabetic patient glucose-insulin regulatory system requires the definition of the *states* of agent's environment, the admissible *actions*, the returned *reward* and the followed *policy*.

State Definition. Glucose-insulin regulatory system represents the environment in our simulation. Different glycemic ranges, which are clinically measured, describe the states of the environment. State 3 is the normoglycemic range of blood glucose concentration which is desired level. Other intervals are unfavorable states which takes the patient in hypo-hyperglycemia. Table 2 demonstrates the states of the environment.

Action Definition. In modeling glucose-insulin regulatory system, insulin infusion rates demonstrate actions. There are five alternative actions for the agent. He obtains the insulin infusion rate by multiplying previous action (a_{t-1}) to a constant number. The five actions we defined here are as follows: The agent can either raise the insulin infusion rate very much

TABLE II
STATES OF THE ENVIRONMENT

State No.	Glycemic Range (mg/dl)	Clinical Description
1	$G \geq 150$	Hyperglycemia
2	$150 \geq G \geq 110$	Slight Hyperglycemia
3	$110 \geq G \geq 70$	Normoglycemia
4	$40 \geq G \geq 70$	Slight Hypolycemia
5	$40 \geq G$	Hypoglycemia

(action 1: $a_{t-1} \times 1.3$), or increase it a bit (action 2: $a_{t-1} \times 1.1$), or even keep the same insulin delivery rate (action 3: a_{t-1}); finally, he can decrease the rate of insulin infusion (action 4: $a_{t-1} \times 0.8$), or decline it a bit (action 5: $a_{t-1} \times 0.95$).

Reward Calculation. The reward is set equal to the difference of the glucose concentration from its target value of 80 mg/dl. This value has been considered as a reference set point in normoglycemic range of blood glucose. The reward $r_t(x,a)$, for taking an action a, from state x, is given by the following equation.

$$r_t(x,a) = -|(G-80)| \quad (8)$$

Where t is the time that agent receives his reward from the environment.

Policy Description. One of the most important features in Q-learning algorithm is the balance between exploration and exploitation. Following a greedy policy constantly, the agent may not visit some states that could turn out to be more profitable. On the other hand, if the agent explores the environment, without exploiting his knowledge, he is not actually learning. Thus, the learning process evolves towards optimal solutions. In this paper, in order to achieve the goal for balance, ε-greedy selection is applied. ε-greedy is a variation of the aforementioned greedy policy. Instead of always taking the best action – as in greedy policy – there is small probability ε, the agent may select another action randomly.

IV. SIMULATION RESULTS

MATLAB is used to simulate the closed-loop system in order to show the validity of the proposed approach.

To verify the physiological model, the controller output $u(t)$ is set to zero and the response of a healthy person and diabetic patient is obtained to show the difference between their glucose regulatory systems. As it can be seen in figure 3, a healthy person's blood glucose value is stabilized in normal value in spite of meal disturbance, but a patient's glucose level remains in much more than basal value.

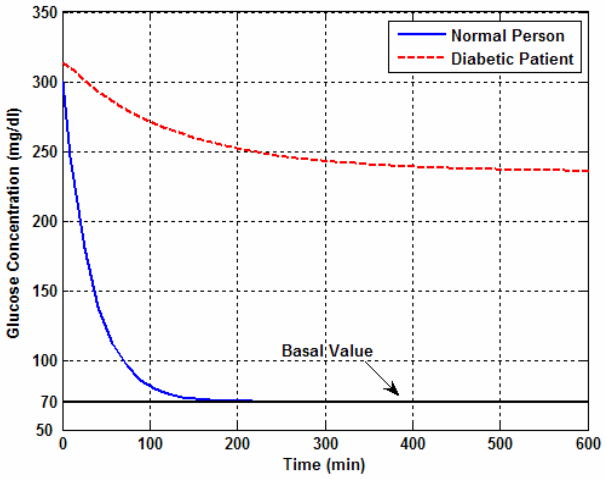

Fig. 3 Healthy person and diabetic patient glucose regulatory system

In second set of simulation the Q-learning algorithm is applied to nonlinear model of the patient. Figure 4 shows the result of simulation in off-line state for 4000 iterations. The following values are chosen for the parameters of the Q-learning algorithm: α=0.5, γ=0.5. The probability ε is set ε=0.5, trying to achieve a good balance between exploration and exploitation. As shown in figure 4, after identifying the environment by the agent, model variables become firm in basal amounts and the Q-value is converges to a constant value. Since the controller implementation requires simple function evaluations, it is much easier to compute than solving an online optimization problem.

In a last set of simulation, fixed Q is applied to the virtual patient and the response of a sick person is examined in an on-line way. Severe initial state of the patient is corrected in spite of meal disturbance. In addition, to check the robustness of the controller to variations in model parameters three sets of parameters for three different patients have been used. It is obvious that the transient responses of different patients in the same controller are different, but in all three cases the plasma glucose and insulin level stabilizes in a reasonable time interval. Figure 5 shows the insulin and glucose profiles and the control function for the same three patients. As it can be seen, optimal insulin delivery rate remains feasible for all the values of uncertain parameters.

The controller performs quite well and keeps the blood glucose level of patient around normal value.

The values that have been used in implementing the model and its parameters are given in Table 3.

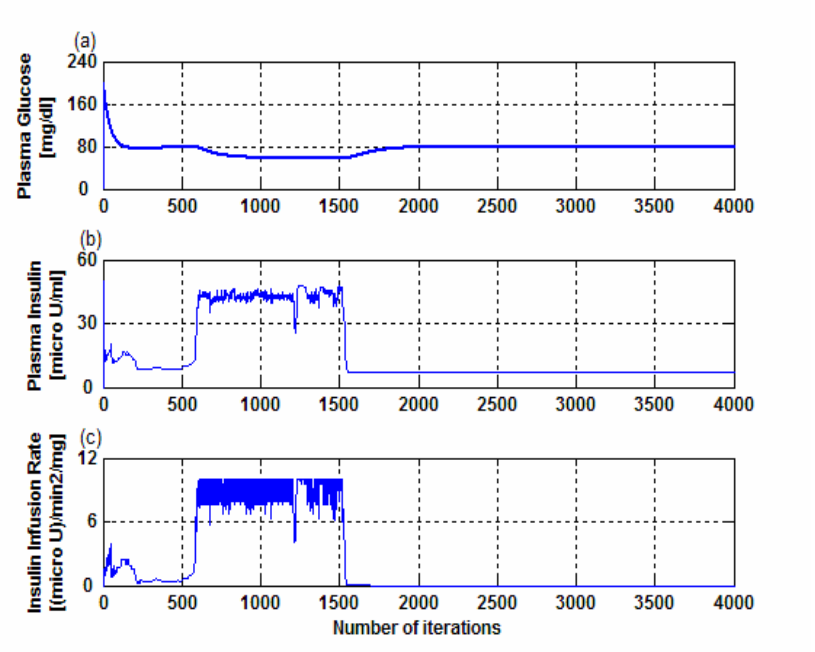

Fig. 4 Glucose-insulin regulatory system with Q-learning algorithm in off-line state. (a) Plasma glucose concentration (b) Plasma insulin concentration (c) insulin infusion rate

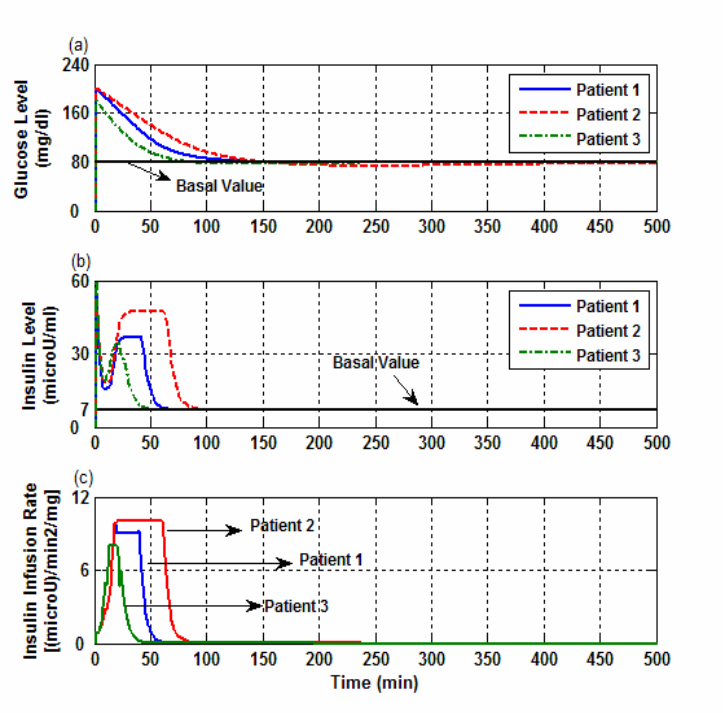

Fig. 5 Closed-loop glucose regulatory system.(a) Plasma glucose concentration. (b) Plasma insulin concentration (c) Exogenous insulin infusion rate.

TABLE III
PARAMETER VALUES

	Normal	Patient 1	Patient 2	Patient 3
p_1	0.0317	0	0	0
p_2	0.0123	0.0107	0.0072	0.0142
p_3	4.92×10^{-6}	5.3×10^{-6}	2.16×10^{-6}	9.94×10^{-6}
γ	0.0039	0.0042	0.0038	0.0046
n	0.2659	0.264	0.2465	0.2814
h	79.0353	80.25	77.5783	82.9370
G_b	80	80	80	80
I_b	7	7	7	7
G_0	291.2	220	200	180
I_0	364.8	50	55	60

V. CONCLUSION REMARKS AND FUTURE WORKS

A. Conclusions

Diabetes management is one of important issues in the field of human regulatory systems, which is discussed in recent years. In This work, a closed-loop control system based on reinforcement learning approach for deriving the explicit insulin delivery rate for type I diabetic patients has been proposed. In order to incorporate knowledge about patient treatment, the controller is designed using Q-learning scheme. It is important to mention that the control algorithm is essence model-free. The proposed controller can successfully tolerate patient variability and dynamic uncertainty while rapidly rejecting meal disturbances and tracking the constant glucose reference. Robustness was tested over a group of three patients with model parameters varying considerably from the averaged model. As shown in this paper, the Q-learning has the potential to synthesize knowledge to treat diseases. Employed control technique reported in this paper is expected to simplify insulin automatic injection mechanism and increase the quality of life, and life expectancy of diabetic patients.

B. Future Works

Future work is conducted to the employing of the dynamic programming analysis approach in designing a robust controller for blood glucose regulation in diabetic patients. It is also considered to compare the result obtained in this paper with some conventional approaches such as PID and H_∞ controllers presented in other researches. In addition, the effect of measurement noise is to be assessed and attenuated. Finally, the inclusion of an exercise regime in the overall model of the Type I diabetic patients in order to have a more realistic simulation will be considered.

REFERENCES

[1] American Diabetes Association, "Standards of medical care for patients with diabetes mellitus," *Diabetes Care*, vol. 26, pp. S33-S50, 2003.
[2] B. Topp, K. Promislow, G. De Vries, "A model of β-cell mass, insulin, and glucose kinetics: Pathway to diabetes," *Journal of Theoretical Biology*. Vol. 206, 2000, pp. 6.5-619.
[3] "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus," *N. Eng. J. Med*, vol. 329, pp. 977-986, 1993, DCCT.
[4] A. Albisser, B. Leibel, T. Ewartz, Z. Davidovac, C. Botz, W. H. Zingg," artificial endocrine pancreas," *Diabetes Care*, vol. 23, 1974, pp. 389-396.
[5] J. Jaremco, O. Rorstad, "Advances toward implantable artificial pancreas for treatment of diabetes," *Diabetes* Care, vol. 21, 1998, pp. 444-450.
[6] M. Nalecz, J. Wojcicki, and I. Zawicki, "Control in artificial pancreas," *IFAC Control Aspects Biomed Eng*, 1987, pp. 123-137.
[7] R. S. parker, G. L. Bowlin, and G. Wnek, Eds. "Insulin delivery," *Encyclopedia of Biomaterials and Biomedical Engineering*. New Yourk: Market Dekker, 2004, pp. 857-866.
[8] F. P. Kennedy, "Recent development in insulin deliver technique: current status or future potential," *Drugs*, vol. 42, 1991, pp. 213-227.
[9] DC. Klonoff. "Continuous glucose monitoring: roadmap for 21st. century diabetes therapy," *Diabetes care* Vol. 28, 2005, pp. 1231-1239.
[10] C. Amaral, B. Wolf, "Current development in non-invasive glucose monitoring," *Journal of Medical Engineering and Physics*, 2007, Elsevier.
[11] TM. Gross, BW bode, D. Einhorn, "Performance evaluation of the minimed continuous glucose monitoring system during patient home use," *Diabetes Technology Theoretical*. vol. 2, 2000, pp. 49-56.
[12] J. Li, Y. Kuang, C. C. Mason, "Modeling the glucose-insulin regulatory system and ultradian insulin secretory oscillations with two explicit time delays," *J. of Theoretical Biology*, vol. 242, 2006, pp. 722-735.
[13] A. De Gaetano and O. Arino, "Mathematical modeling of the intravenous glucose tolerance test", *J. Math. Biol.* Vol. 40, 2000, pp. 136-168.
[14] R. N. Bergman, L. Philips, and C. Cobelli, "Physiological evaluation of the factors controlling glucose tolerance in man," *Journal of Clinical Investigation*, vol. 68, 1981, pp. 1456-1467.
[15] J. T. Sorenson, "A physiological model of glucose metabolism in mans its use to design and assess improved insulin therapies for diabetes," ", Ph.D. Dissertation, Chem Eng. Dep., Massachusets Inst. Technol, Cambridge, 1985
[16] Ch. Li and R. Hu, "Simulation study on blood glucose control in diabetics", *Proc. IEEE Int. Conf. on Biomed. and Bioinf Eng*, 2007, pp. 1103-1106.
[17] F. Chee, T. L. Fernando, A. V. Savkin, and V. Heeden, "Expert PID control system for blood glucose control on critically ill patients," *IEEE Trans. Biomed. Eng.*, vol. 7, No. 4, 2003, pp. 419-425.
[18] Z. H. Lam, J. Y. Hwang, J. G. Lee, J. G. Chase, and G. C. Wake, "Active insulin infusion using optimal and derivative-weighted control," *Medical Engineering & Physics.* vol. 24, 2002, pp. 663-672.
[19] K. H. Kientiz and T. Yoneyame, "A robust controller for insulin pupms based on H-infinity theory" *IEEE Transaction on Biomedical Eng.*, vol. 40, 1993, pp. 1133-1137.
[20] Sh. Yasini, A. Karimpour, M. B. Naghibi-Sistani, S. Ghareh., "An automatic insulin infusion system based on H-infinity control technique," Procceding of the 2008 IEEE, CIBEC'08, 2008, pp. 1-5, Cairo, Egypt.
[21] F. Chee, AV. Savkin, TL Fernando, S. Nahavandi, "optimal H-infinity insulin injection control for blood glucose regulation in diabetic patients," vol. 52, No. 10, 2005, pp. 1625-1631.
[22] M. S. Ibbini, M. A. Masadeh, and M. M. Bani Amer, "A semi closed loop optimal control system blood glucose level in diabetics", *J. Medical Eng. & Tech.* Vol. 28., 2004, pp. 189-196.
[23] M. E. Fisher, "A semi closed-loop algorithm for control of blood glucose levels in diabetics", *IEEE Trans. On Biomed. Eng*, vol.38, 38. No. 1, 1991.
[24] R. S. Sutton, A. G. Barto, *Reinforcement learning: An introduction*, 1989, MIT Press.
[25] R. N. Bergman, D. T. Finegood, M. Ader, "Assessment of insulin sensitivity invivo", *Endocrine Reviews*, Vol. 6, No. 1, 1985, pp. 45-85.

[26] C. Neatpisarnvanit, JR Boston, "Estimation of plasma insulin for plasma glucose. IEEE Transaction on Biomedical Engineering," vol. 49, No. 11, 2002, pp.1253-1259.
[27] C. Watkins, *Learning from delayed rewards*, Ph.D. Dissertation Cambridge University, 1998.