Artificial Intelligence in Medicine (2004) **32**, 171–181

ELSEVIER

ARTIFICIAL INTELLIGENCE IN MEDICINE

http://www.intl.elsevierhealth.com/journals/aiim

Evaluation of glucose controllers in virtual environment: methodology and sample application

Ludovic J. Chassin, Malgorzata E. Wilinska, Roman Hovorka*

Diabetes Modelling Group, Department of Paediatrics, University of Cambridge, Box 116, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK

Received 23 May 2002; received in revised form 6 July 2003; accepted 27 February 2004

KEYWORDS
Simulation environment;
Adaptive systems;
Closed-loop control;
Type 1 diabetes mellitus

Summary

Objective: Adaptive systems to deliver medical treatment in humans are safety-critical systems and require particular care in both the testing and the evaluation phase, which are time-consuming, costly, and confounded by ethical issues. The objective of the present work is to develop a methodology to test glucose controllers of an artificial pancreas in a simulated (virtual) environment.
Material and methods: A virtual environment comprising a model of the carbohydrate metabolism and models of the insulin pump and the glucose sensor is employed to simulate individual glucose excursions in subjects with type 1 diabetes. The performance of the control algorithm within the virtual environment is evaluated by considering treatment and operational scenarios.
Results: The developed methodology includes two dimensions: testing in relation to specific life style conditions, i.e. fasting, post-prandial, and life style (metabolic) disturbances; and testing in relation to various operating conditions, i.e. expected operating conditions, adverse operating conditions, and system failure. We define safety and efficacy criteria and describe the measures to be taken prior to clinical testing. The use of the methodology is exemplified by tuning and evaluating a model predictive glucose controller being developed for a wearable artificial pancreas focused on fasting conditions.
Conclusion: Our methodology to test glucose controllers in a virtual environment is instrumental in anticipating the results of real clinical tests for different physiological conditions and for different operating conditions. The thorough testing in the virtual environment reduces costs and speeds up the development process.

1. Introduction

The identifying property of adaptive systems is their ability to perform in a changing environment. This can be extended to another two levels to include

* Corresponding author. Tel.: +44 1223 762 862; fax: +44 1223 336 996.
E-mail address: r.hovorka@uk.avecho.com (R. Hovorka).

0933-3657/$ – see front matter
doi:10.1016/j.artmed.2004.02.006

adaptation to a similar setting and adaptation to a new/unknown application as specified by the EUNITE network of excellence on smart adaptive systems [1].

Adaptive systems to deliver medical treatment in humans are safety-critical systems and require particular care in both the testing and the evaluation phase. Examples of such systems include control in anaesthesia [2] and insulin treatment for type 1 diabetes mellitus [3].

The development, evaluation, and testing are time-consuming, costly, and confounded by ethical issues. The system has to be evaluated in a range of treatment scenarios to guarantee safety and to demonstrate adaptivity leading to acceptable treatment efficacy. This is a complicated process and only the first step in the development of commercially viable medical applications of adaptive technology.

Simulation (virtual) environments offer reductions in human, time, and financial costs. Simulated clinical tests have become an essential part of the drug development process [4] but have yet to find their way into the development of medical devices beyond the trivial testing on a representative patient, although some more elaborate examples can be found in areas such as the control in anaesthesia [5,6]. The rationale is simple and appealing. Instead of conducting trials on real subjects, a collection of virtual subjects is tested in virtual computer space that offers a close-to-real-life behaviourally rich environment. The approach can be utilised for both system development and system evaluation.

Type 1 diabetes is characterised by an absolute insulin deficiency. Current treatment relies on insulin delivery by multiple daily insulin injections or a continuous subcutaneous insulin infusion (CSII). Recent technological advances in minimally invasive continuous glucose measurement techniques fuel research into the development of a technologically based artificial pancreas consisting of a glucose sensor, a glucose controller, and an insulin pump. Although the availability of a reliable, longer term glucose sensor remains the rate limiting factor, the development of a glucose controller is difficult. The system is subjected to long time delays originating, for example, in the absorption of subcutaneously injected insulin and the prolonged duration of insulin action, large inter-subject variability in insulin requirements and also changes in insulin need throughout the day. It has been shown that an adaptive control strategy is required [7].

Diabetes simulators have already established themselves as important educational tools [8,9]. They have been employed to gain insights into pathophysiological conditions such as the development of insulin resistance in type 2 diabetes [10] and have found place in the development and performance assessment of control algorithms for automated insulin release control system [11,12]. But, to our knowledge, a complete and systematic methodology to test glucose controllers has yet to be developed.

The objective of the present paper is to describe a methodology to test glucose controllers in a virtual environment. The methodology includes two dimensions, testing in relation to specific life style conditions (for example, fasting and post-prandial), and testing in relation to operating conditions such as expected or adverse operating conditions. We define safety and efficacy criteria and describe the measures to be taken prior to clinical testing. The use of the methodology is exemplified by tuning and evaluating a model predictive glucose controller in the process of being developed for a wearable artificial pancreas [13] focusing on testing under fasting conditions.

The testing is divided into the two aforementioned dimensions as this provides a coherent framework to carry out systematic evaluation proceeding from simple (most common) to more complex (less common) scenarios under which the system is expected to perform.

The paper briefly describes the simulation environment and the model predictive controller. The focus is on the methodology for testing glucose controllers, which is described in detail together with its sample application.

2. Methods

2.1. Simulation environment

The simulation environment represents the interaction between a virtual subject with type 1 diabetes, a measurement model, a glucose controller, and an insulin pump, see Fig. 1. The simulation environment is implemented in Matlab and Simulink.

The measurement model represents the measurement process and includes the properties of the glucose sensor such as the measurement error. The insulin pump delivers the insulin into the subcutaneous tissue of the virtual subject and can also be parameterised to include an error in the insulin delivery due to technical limitations and tissue properties.

At present, the environment represents 18 virtual subjects with type 1 diabetes. Differentiation

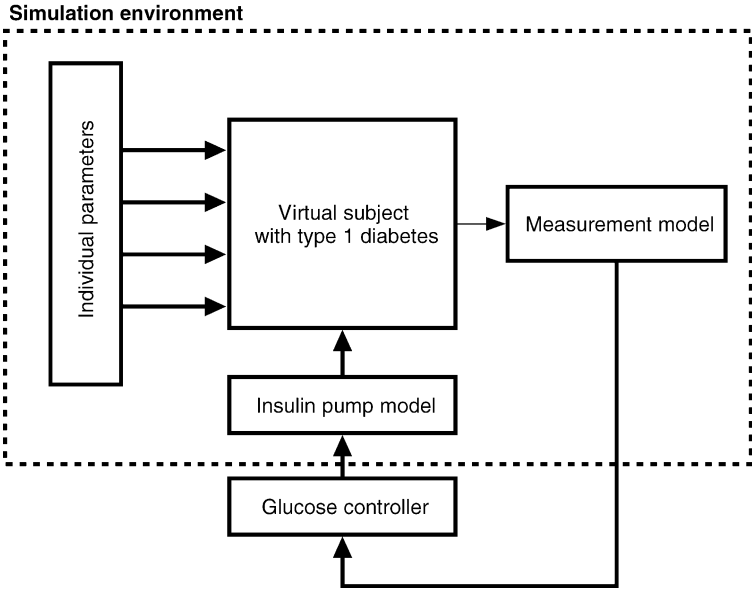

Figure 1 Overview of a simulation (virtual) environment. All components except the glucose controller(s) are implemented within the environment facilitating tuning and comprehensive testing of the controller.

among virtual subjects is achieved by assigning a unique set of parameters to the glucoregulatory model. The parameters were obtained by analysing data collected overnight in real subjects with type 1 diabetes during variable intravenous insulin infusion and 30-min plasma glucose (PG) sampling. Parameters not obtainable from these data, such as gut absorption or insulin absorption from the subcutaneous tissue, were drawn from informed (population) probability distributions.

Intra-individual variability of the glucoregulatory system is implemented in the virtual patients by superimposing sinusoidal oscillations on model parameters. These include fast oscillations (3-h period) and slow oscillations (24-h period) in combination with a low amplitude (5% of parameter value) and a high amplitude (30%). The phase is also drawn from a random (uniform) distribution for each parameter and this provides for virtual subjects with widely ranging behaviour characteristics.

The virtual subject is based on a physiologically based compartment glucoregulatory model described by a set of first-order differential equations following work by Hovorka et al. [14].

The glucoregulatory model represents input–output relationship between subcutaneous insulin infusion on input and intravenous glucose concentration on output. Meal ingestion and intravenous glucose infusion represent additional inputs. The core of the model is outlined in Fig. 2.

The model consists of a glucose subsystem (glucose absorption, distribution, and disposal), an insulin subsystem (insulin absorption, distribution, disposal), and an insulin action subsystem (insulin action on glucose transport, disposal, and endogenous production).

To exemplify a subsystem, glucose kinetics is described by a set of differential equations:

$$\frac{\mathrm{d}Q_1(t)}{\mathrm{d}t} = -\left[\frac{F_{01}^{c}}{V_G G(t)} + x_1(t)\right] Q_1(t) + k_{12} Q_2(t) - F_R + U_G(t) + \mathrm{EGP}_0[1 - x_3(t)]$$

$$\frac{\mathrm{d}Q_2(t)}{\mathrm{d}t} = x_1(t) Q_1(t) - [k_{12} + x_2(t)] Q_2(t)$$

$$y(t) = G(t) = \frac{Q_1(t)}{V_G}$$

where Q_1 and Q_2 represent the masses of glucose in the accessible (where measurements are made) and non-accessible compartments, k_{12} represents the transfer rate constant from the non-accessible to the accessible compartment, V_G represents the distribution volume of the accessible compartment, y and G the (measurable) glucose concentration, EGP_0 represents endogenous glucose production (EGP) extrapolated to the zero insulin concentration, F_{01}^{c} the total non-insulin-dependent glucose flux corrected for the ambient glucose concentration, and F_R the renal glucose clearance above the glucose threshold of 9 mmol/l.

Similar but simpler differential equations describe gut absorption, insulin absorption from subcutaneous depot, and insulin action.

The validity of the core of the model, i.e. the model of glucose kinetics and insulin action, has been demonstrated in healthy subjects during intravenous glucose tolerance test (IVGTT) [14]. The

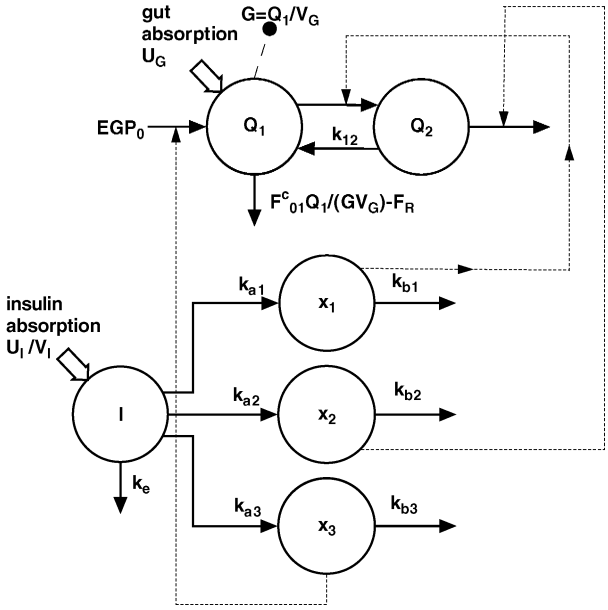

Figure 2 The core of the virtual subject with type 1 diabetes contains a compartment model of glucose—insulin system. Q_1 and Q_2 represent glucose masses in plasma and non-accessible compartments, I represent plasma insulin, x_i represent insulin action on glucose transport, disposal, and endogenous glucose production.

limitation of the simulation environment is that an extrapolation of the parameters from healthy to type 1 diabetes subjects was required. Another limitation is that the suppression of hepatic glucose production was not studied during meal when insulin and glucose is elevated for a considerable longer period than during IVGTT.

We have evaluated the insulin absorption submodel and the insulin kinetics submodel in nine subjects with type 1 diabetes during continuous insulin infusion (data not shown). The submodels were found adequate although a marginal improvement was obtained by postulating a different absorption rate for the short-acting insulin analogue administered via continuous infusion (a slightly faster absorption rate) compared to the prandial bolus (a slower absorption rate). The primary source for the time-to-maximum of insulin absorption comes from two studies assessing the kinetics of short-acting insulin after bolus administration [15,16].

Probably, the weakest part of the model in terms of experimental validity is the gut absorption submodel. Detailed investigations of the gut absorption rate usually employ the ingestion of oral glucose or liquid meal [17]. In the post-prandial state, appearance of glucose resulting from meal/glucose ingestion is up to three times higher than fasting endogenous production [17], which is simultaneously suppressed by more than half after meal [18] indicating the dominating influence of exogenous appearance. Glucose absorption is difficult to quantify in in vivo conditions. It has a high intra-individual variability [19] and a comprehensible mathematical description relating the absorption rate to food composition and the meal size has yet to be established despite considerable extended effort and advances in the field [18,20—24].

2.2. Model predictive controller of glucose concentration

The model predictive control (MPC) [25] is an emerging methodology to facilitate control of systems with long time delays and open loop characteristics. When combined with adaptive capabilities, it promises to tackle successfully problems such as the control of glucose concentrations.

In this study, we evaluate an MPC-based glucose controller [13]. The controller has been designed to be used with an intravenous and subcutaneous glucose measurement, in combination with a subcutaneous insulin infusion of short acting insulin such as Lispro.

Model predictive control normally relies on the internal ''model'' representation of the underlying system adopting linearised version [26] of one of the many models of the glucoregulatory system [8,27—29]. However, our controller retains the full non-linear model to minimise the model mis-specification error.

The controller adopts a model with an identical structure to that used by the virtual subject with type 1 diabetes. Significantly, the underlying model is parameter-rich and it is not possible to estimate exhaustively model parameters from the observed input (insulin infusion) — output (plasma glucose) relationship. For the purposes of the controller, model quantities are divided into model constants and model parameters with the objective to reduce the number of parameters. Model constants are fixed and represent those quantities which (i) are not a priori identifiable [30] or (ii) are unlikely to be identifiable from the data (a posteriori non-identifiability). Only the remaining seven parameters are estimated from the data, i.e. non-insulin dependant glucose flux, EGP extrapolated to zero insulin concentration, insulin sensitivity of distribution/transport, insulin sensitivity of EGP, time-to-maximum of absorption of subcutaneously injected short-acting insulin, CHO availability, and time-to-maximum of CHO absorption.

The MPC controller includes the following components: parameter optimiser, target projector, dose optimiser, and safety schemes.

The parameter optimiser employs the maximum a posteriori Bayesian parameter estimation

technique (a multivariate log prior distribution for the seven parameters is assumed). Glucose measurements from a ''learning window'', i.e. a time period immediately preceding the control time, are employed in the parameter estimation. Three lengths of the learning window are predefined, short, medium, and long, to be able to deal with both a time-invariant (or slowly varying) underlying system, which is best identified over a long learning window, and a time-variant system best identified over a shorter learning window. The target projector calculates target trajectory, i.e. the desired glucose profile. The dose optimiser adopts non-linear function minimisation as the underlying model is non-linear. The control action is the first insulin infusion rate of a control sequence, which is calculated by minimising an objective function with two components: (i) the adherence of the predicted glucose to the target trajectory and (ii) the variation in the control sequence implemented as the norm of first differences in the sequence of insulin infusion rates.

The ''aggressiveness'' is a parameter of the MPC controller, which quantifies the penalty associated with the variation in the insulin infusion rate. A large value of aggressiveness increases the risk of hypoglycaemia whereas a low level results in a too slow normalisation of glucose levels.

2.3. System tuning

The simulation environment represents an ideal tool to tune the controller and also to investigate the effect of various system set-ups. We exemplify this by finding an optimal value for the aggressiveness parameter of the controller and by assessing the effect of constraints in insulin infusion particularly the maximum infusion rate, see Fig. 3.

2.4. Methodology to test blood glucose controllers employing simulator of glucose metabolism

The methodology evaluates the glucose controller in three (simulated) *physiological* conditions: fasting, post-prandial, and life style (metabolic) disturbances such as exercise, see Fig. 4.

In each physiological condition, we are concerned with the performance of the glucose controller under three *operating* conditions, expected operating conditions (EOC), adverse operating conditions, and system failure, see Fig. 5.

2.4.1. Expected operating conditions

EOC represents the typical conditions in which the controller is expected to work. These include the expected properties of the sensing device, properties of the target population, and properties of the insulin delivery device.

The glucose sensor is characterised by its measurement error and a measurement delay associated with subcutaneous glucose sensing (the delay is due to, for example, the kinetic properties of the glucose transport between plasma and the interstitial fluid).

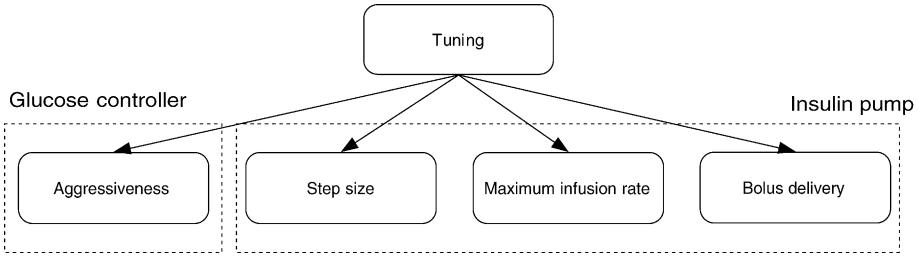

Figure 3 Simulation environment is instrumental in tuning the controller and assisting in the evaluation of settings of other system components such as the insulin pump.

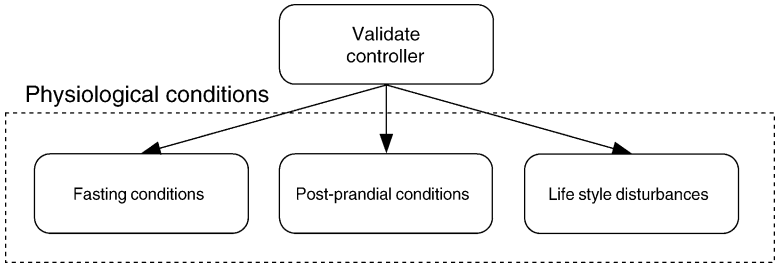

Figure 4 Glucose controller must be evaluated under different physiological conditions.

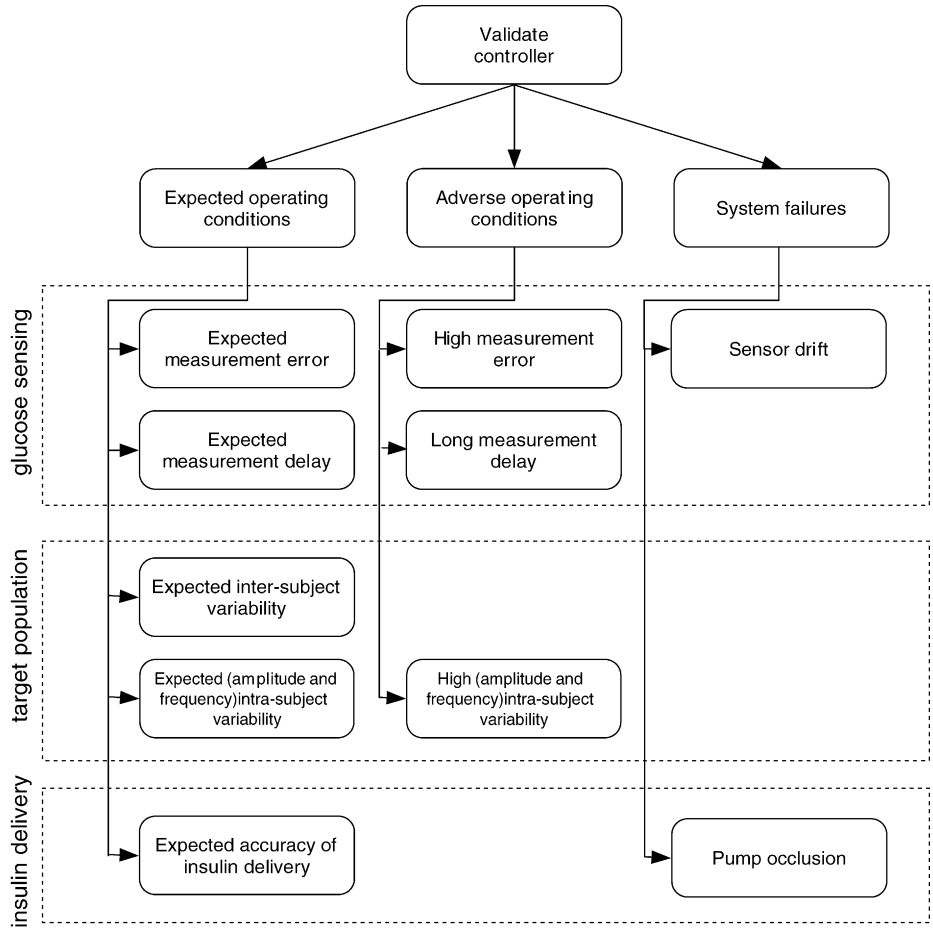

Figure 5 Overview of the validation methodology for blood glucose controllers (see text for details).

The target population is characterised by its inter-subject and intra-subject variability. Subjects differ in how much insulin they need (inter-subject variability) which reflects their insulin sensitivity and other parameters of the glucoregulatory system. Second, insulin needs vary throughout the day due to, for example, diurnal changes in insulin sensitivity (intra-subject variability).

The insulin pump delivers insulin according to its technical specification (a step-by-step motor) and subjects to constraints such as the maximum delivery rate and the step resolution.

The set-up for early clinical testing includes intravenous glucose sampling and subcutaneous insulin delivery. We expect a 3—5% coefficient of variation (CV; absolute measurement error divided by ''true'' value $\times$ 100%) of the measurement error, a 3-min measurement delay (representing time needed to analyse plasma glucose on a bed-side analyser), a intra-subject variability (oscillations) of the parameters of the glucoregulatory system with a 5% CV and a 3 or 24-h period, a 3-min step delivery of insulin with a resolution 0.1 U/h and a maximum infusion rate 4 U/h, and a 15% CV of the error associated with insulin delivery.

The expected inter-subject variability of parameters of the glucoregulatory system is given by the parameters of the virtual subjects derived from real data.

2.4.2. Adverse operating conditions

Properties of the system components/environment can temporarily deteriorate and it is essential to assess the performance of glucose controllers under these adverse conditions.

In particular, the measurement error associated with the glucose sensor can increase or the measurement delay can be prolonged. This also includes an increased amplitude of the intra-subject variability of the parameters of the glucoregulatory system.

Our implementation of the adverse operating conditions increased the CV of measurement error to 8 and 15%, the measurement delay to 40 min, and the amplitude of the intra-subject variability to 30% with an unaltered period of 3 or 24 h. The measurement delay in the adverse operating conditions assumes subcutaneous glucose sampling via extracorporal open flow microperfusion [31] and consists of a 10-min physiological delay, a 20-min dead space

of the sampling cannula, and an additional adverse-conditions-related 10-min delay.

2.4.3. System failure

System failure may have catastrophic consequences. Glucose controllers need to have a built-in safety (protection) scheme to deal with exceptional circumstances [32] with the objective of minimising the risk of insulin overdose and subsequent hypoglycaemia.

The tests we propose to execute are related to a sensor drift and to an insulin pump occlusion. The sensor drift represents the situation where there is a progressively increasing bias between glucose measurements and the actual glucose levels and if the sensor overestimates actual glucose levels this may lead to insulin overdosing. A pump/tissue occlusion lasting several hours may lead to the creation of an insulin reservoir in the pump/tissue, which can abruptly enter the system. This problem can be exacerbated by the glucose controller increasing the insulin infusion during the occlusion period to ''counteract'' increasing glucose concentration.

2.4.4. Assessing glucose control

The glucose control is evaluated using two criteria. The primary criterion is the *safety* of glucose control. The secondary criterion is the *efficacy* of glucose control. The safety criterion is concerned with the avoidance of low plasma glucose (hypoglycaemia), the efficacy criterion with the avoidance of high plasma glucose (hyperglycaemia). The rationale behind the definitions of the two criteria is that of an acute danger such as unconsciousness or coma associated with hypoglycaemia, and that of the development and progression of microvascular and macrovascular complications of diabetes (limb amputations, blindness, renal failure, etc.), which is associated with a prolonged elevation of plasma glucose. The criteria are defined in Table 1.

2.4.5. Safety criterion

The safety criterion assesses the safety of the glucose control. This is achieved by evaluating the number of virtual subjects who present hypoglycaemia (2.0 mmol/l $<$ PG $\leq$ 3.3 mmol/l) and serious hypoglycaemia (PG $\leq$ 2.0 mmol/l) during a relatively short duration ($\leq$24 h) of virtual tests.

''Excellent'' safety is achieved by avoiding hypoglycaemia and severe hypoglycaemia in all subjects. ''Good'' safety is characterised by a very low risk of hypoglycaemia ($\leq$5% of subjects with hypoglycaemia) without the presence of severe hypoglycaemia. ''Satisfactory'' safety is aimed to reflect incidence of hypoglycaemia with CSII (the current ''gold'' standard); there are limited comparable data and we set the limits ($\leq$20% of subjects with hypoglycaemia; no severe hypoglycaemia) more stringently than those reported in a study by Renner et al [33] (about 40% subjects experienced hypoglycaemia in a day). ''Unsatisfactory'' safety is characterised by levels of hypoglycaemia in more than 20% subjects or by the presence of severe hypoglycaemia in one or more subjects.

The safety criterion is identical for all three ''physiological'' conditions, i.e. fasting, post-prandial, and metabolic disturbances.

Table 1 Definition of safety and efficacy criteria for glucose controllers

Grade	Fasting conditions	Post-prandial state	Metabolic disturbances such as physical exercise, alcohol intake, etc.
Safety			
Excellent	No hypo[a] and no serious hypo[b]	As for fasting conditions	As for fasting conditions
Good	$\leq$ 5% subjects with hypo and no subject with serious hypo	As for fasting conditions	As for fasting conditions
Satisfactory	$\leq$ 20% subjects with hypo and no subject with serious hypo	As for fasting conditions	As for fasting conditions
Unsatisfactory	$>$ 20% subjects with hypo or at least one subject with serious hypo	As for fasting conditions	As for fasting conditions
Efficacy			
Excellent	PG $\leq$ 6 mmol/l	2 h PG $\leq$ 8 mmol/l	As for fasting conditions
Good	PG $\leq$ 7 mmol/l	2 h PG $\leq$ 9 mmol/l	As for fasting conditions
Satisfactory	PG $\leq$ 8 mmol/l	2 h PG $\leq$ 11 mmol/l	As for fasting conditions
Unsatisfactory	PG $>$ 8 mmol/l	2 h PG $>$ 11 mmol/l	As for fasting conditions

[a] hypo is defined as 2.0 mmol/l $<$ PG $\leq$ 3.3 mmol/l.
[b] serious hypo as PG $\leq$ 2.0 mmol/l.

It is important to stress that simulated clinical trials are run for 8–24 h and thus the safety criterion is defined taking into account the relatively short duration of the trials quantifying percentage of subjects with hypoglycaemia rather than incidence of hypoglycaemia per patient year as is standard in longitudinal studies such as DCCT [33].

2.4.6. Efficacy criterion

The efficacy criterion evaluates the ability of glucose controllers to achieve glucose levels, which are known to avoid diabetes complications.

In the fasting state, ''excellent'' efficacy gives glucose levels ($\leq$6 mmol/l) similar to those observed in healthy subjects. ''Good'' efficacy (PG $\leq$ 7 mmol/l) represents the WHO limit for the diagnosis of impaired glucose tolerance [34]. The ''satisfactory'' efficacy (PG $\leq$ 8 mmol/l) corresponds approximately to the best currently available management practice as documented by intensive insulin therapy in DCCT [33]. ''Unsatisfactory'' efficacy is that which is less satisfactory, i.e. worse than the current best management practice.

In the post-prandial state, we focus on 2-h post-prandial glucose, which is widely used for the diagnosis of diabetes and in the assessment of glucose control. ''Excellent'' efficacy (2 h PG $\leq$ 8 mmol/l) corresponds to the WHO diagnostic limit of diabetes [34]. ''Satisfactory'' efficacy (2 h PG $\leq$ 11 mmol/l) represents a slightly worse control than that observed with the best currently available management practice as exemplified by the intensive therapy during DCCT [35]. ''Unsatisfactory'' efficacy represents control worse than what is currently available.

In the conditions of life style disturbances, the efficacy criterion is the same as during the fasting conditions.

3. Results and discussion

3.1. Simulation environment

A sample outcome of a simulated clinical trial in a virtual subject is shown in Fig. 6. The test included the digestion of a meal with co-administration of a ''manually'' determined insulin bolus. The closed-loop control commenced 3 h after the meal.

3.2. System tuning

The results of the system tuning are shown in Table 2. The results refer to testing during fasting conditions.

In the first instance we investigated the effect of ''aggressiveness'' on the performance of the controller. The best trade-off was with aggressiveness 5 (unitless) which gave good safety and excellent efficacy. Higher values of aggressiveness maintained excellent efficacy but safety was compromised. Lower values of aggressiveness reduced efficacy.

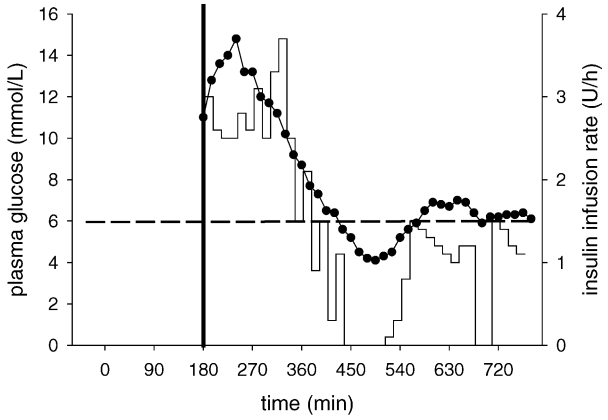

Figure 6 Example of glucose control using simulator. At time 0 min, a synthetic type 1 diabetes subject had a meal and received insulin bolus. At 180 min (vertical bar), glucose was artificially raised to 11 mmol/l and the subcutaneous insulin infusion was initiated. A model based glucose controller calculated the infusion rate every 15 min using intravenous glucose measurements also taken every 15 min.

All future runs with the controller were therefore run with the aggressiveness set to 5 (unitless).

A close inspection of the data collected during the ''aggressiveness'' tests indicated that some virtual subjects were temporarily infused with insulin rates above the upper limit currently employed during CSII. Clinical considerations suggested employing an upper limit of 4 U/h and technical considerations of an insulin pump to be employed in real clinical trials (to facilitate a fast manual change of the insulin pump rate) implied that a 2 U/h limit

Table 2 Tuning controller and insulin pump settings

	Safety	Efficacy
Aggressiveness		
1	Good	Good
3	Good	Good
5	Good	Excellent
7	Satisfactory	Excellent
1000	Satisfactory	Excellent
Maximum infusion rate 4 U/h with step size 0.2 U/h	Good	Excellent
Maximum infusion rate 2 U/h with step size 0.1 U/h	Excellent	Good
Bolus delivery	Excellent	Good

should also be considered. The former limit was tested with a reduced step size of 0.2 U/h (in all other tests a step size resolution was 0.1 U/h) again considering a fast manual alteration of the pump rate.

The limit of 4 U/h lead to the same results as the limit-free testing. The 2 U/h limit improved safety. However, efficacy was reduced due to several insulin resistant subjects receiving too little insulin. We decided to adopt the 4 U/h and not the 2 U/h limit for future testing. The step size was set at 0.1 U/h as this was subsequently found technically feasible.

Insulin pumps deliver insulin in microboluses 0.5–3 min apart. This quantum delivery may introduce a discrepancy between the expected and truly infused insulin at times when infusion rates change considerably. The discrepancy can be avoided by administering insulin not as a continuous insulin infusion but as boluses given at the start of the control cycle (in our system, insulin rate changes every 15 min). The performance of bolus delivery was evaluated and gave excellent safety and good efficacy. The efficacy just failed to reach the excellent grade. Thus, in purely performance terms, bolus delivery of insulin every 15 min appears to be a preferable mode of insulin delivery. We did not, however, retain this feature as we considered that a zero fluid flow for 15 min at the tip of the cannula might increase the risk of occlusion.

3.3. Application of the methodology to test blood glucose controllers on simulator

To exemplify the use of the methodology we evaluated the performance of the model predictive controller under fasting conditions. The results are shown in Table 3.

Under expected operating conditions, the controller achieved excellent safety and good efficacy. No hypoglycaemia events were observed. This provided the reassurance for a subsequent real clinical testing, which gave similar results to those obtained in the simulation study [36]. The similarity of the two sets of results provides support for the validity of the virtual environment.

Under adverse operating conditions, the high measurement error reduced safety while maintaining efficacy. The reduction in safety can be at least in part attributed to ''false'' hypoglycaemia events, i.e. plasma glucose above the hypoglycaemia threshold can be ''measured'' below the threshold after adding the measurement error and this situation will be exacerbated with an increased measurement error. The long measurement delay reduced efficacy but still achieved the satisfactory grade. The high intra-subject variability was the only set of conditions with unsatisfactory safety. This was due to the presence of one severe hypoglycaemia (out of 36 virtual tests). This indicates that large variations (30% of nominal values) in individual parameters are difficult to deal with by the glucose controller.

Of the two types of system failure, only pump occlusion was tested. The 2-h occlusion gave satisfactory safety and good efficacy as assessed by post-occlusion plasma glucose over a period of 5 h. This was an important observation as pump occlusion is not a rare event. Modern insulin pumps are equipped with pressure sensors to detect the occlusion but an additional safety margin is beneficial. The sensor drift has not been tested, as limited information is known about the glucose sensor being developed for this particular version of the artificial pancreas.

3.4. General discussion

Development and testing in a simulation environment is an appealing and worthwhile strategy to lower costs and to reduce the development time. The validity of the approach lies in the creation of a virtual environment faithfully representing real world conditions. In the medical field, the richness

Table 3 Summary of validation results for model predictive controller during fasting conditions

	Fasting conditions	
	Safety	Efficacy
Expected operating conditions	Excellent	Good
Adverse operating conditions		
High measurement error	Satisfactory	Good
Long measurement delay	Excellent	Satisfactory
High intra-subject variability	Unsatisfactory	Excellent
System failure		
Sensor drift	ND[a]	ND
Pump occlusion	Satisfactory	Good

[a] Not determined.

of the environment has to reflect the variability between individuals, variability within individuals, and characteristics of the technical components such as measurement error.

In this paper we are concerned with the development of a testing methodology for glucose controllers. Several notes related to the development and validation of the virtual environment apply. First, the virtual development should be based on real data. It is clear, however, that there will be components for which our knowledge is incomplete and/or inconsistent. This requires qualified guesses to be made about underlying probability distributions of some parameters and properties. Second, the validation of the virtual environment is an important issue but this can normally be resolved only after real clinical tests have been conducted.

The methodology presented in this paper has general applicability for other medical applications of adaptive/hybrid techniques and could be used for other types of clinical adaptive systems and beyond a particular type of a controller.

The evaluation is carried out with respect to external and internal dimensions. The former reflects that adaptive systems work within a behaviour-rich environment. Normally, a set of predefined scenarios could be specified, e.g. *life style conditions* could be defined, and testing should proceed from simple to more complex (or most common to less common) scenarios. The latter internal dimension of the evaluation is related to settings of *physiological and technical parameters* of the system (stressing that the human is part of system) proceeding from the normal (expected) operating conditions to less favourable until those classified as system failure.

An important part of the methodology is the definition of evaluative criteria. In the medical domain these neatly divide into the assessment of *safety* (is the system safe to use) and *efficacy* (does the system achieves its primary goal) following the approach paved by the drug development field. The safety criteria are likely to be shared when evaluating the internal and external dimensions. The efficacy criteria are likely to differ among various scenarios reflecting higher expectations on the system performance under standard/most prevalent conditions and less demanding expectations under less favourable conditions.

In conclusion, we have developed a methodology to test glucose controllers in a simulated (virtual) environment. The methodology is instrumental in anticipating the results of real clinical tests for different physiological conditions and for different operating conditions. The thorough testing in the virtual environment reduces costs and speeds up the development process.

Acknowledgements

This work was supported by the 5th EC Framework Programme under Grant No. IST-1999-14027 (ADICOL).

References

[1] Anguita D. Smart adaptive systems: State of the art and future directions of research. http://www.eunite.org/eunite/events/eunite2001/Papers/13340_P_Anguita.pdf (accessed 17 March 2003).
[2] Shieh JS, Linkens DA, Peacock JE. Hierarchical rule-based and self-organizing fuzzy logic control for depth of anaesthesia. IEEE Trans Syst Man Cyb 1999;29:98—109.
[3] Shichiri M, Kawamori R, Hakui N, Yamasaki Y, Abe H. Closed-loop glycemic control with a wearable artificial endocrine pancreas. Variations in daily insulin requirements to glycemic response. Diabetes 1984;33:1200—2.
[4] Sheiner LB, Hashimoto Y, Beal SL. A simulation study comparing designs for dose ranging. Stat Med 1991;10:303—21.
[5] Mahfouf M, Abbod MF, Linkens DA. The design of supervisory rule-based control in the operating theatre via an anaesthesia simulator. Expert Syst 2002;19:11—20.
[6] Mason DG, Linkens DA, Edwards ND, Reilly CS. Development of a portable closed-loop atracurium infusion system: Systems methodology and safety issues. Int J Clin Monit Com 1996;13:243—52.
[7] Fischer U, Schenk W, Salzsieder E, Albrecht G, Abel P, Freyse EJ. Does physiological blood glucose control require an adaptive control strategy? IEEE Trans Biomed Eng 1987;34:575—82.
[8] Berger M, Rodbard D. Computer-simulation of plasma-insulin and glucose dynamics after subcutaneous insulin injection. Diabetes Care 1989;12:725—36.
[9] Lehmann ED. Preliminary experience with the Internet release of AIDA — an interactive educational diabetes simulator. Comput Meth Programs Biomed 1998;56:109—32.
[10] Summers RL, Montani JP, Woodward LH, Coleman TG, Hall JE. Theoretical analysis of the mechanisms of chronic hyperinsulinemia. Comput Biol Med 1997;27:249—56.
[11] Asakawa N, Saito Y, Yamasaki Y, Kawamori R, Shichiri M. Validation of closed-loop subcutaneous insulin infusion algorithm-application of subcutaneous insulin absorption kinetics. Diabetes Res 1987;5:193—8.
[12] Trajanoski Z, Regittnig W, Wach P. Simulation studies on neural predictive control of glucose using the subcutaneous route. Comput Meth Programs Biomed 1998;56:133—9.
[13] Hovorka R, Canonico V, Chassin LJ, Haueter U, Massi-Benedetti M, Orsini-Federici M, et al. Non-linear model predictive control of glucose concentration in subjects with type 1 diabetes. Physiol Meas 2004;25:905—20.
[14] Hovorka R, Shojaee-Moradie F, Carroll PV, Chassin LJ, Gowrie IJ, Jackson NC, et al. Partitioning glucose distribution/transport, disposal, and endogenous production during IVGTT. Am J Physiol 2002;282:E992—1007.

[15] Rave K, Heinemann L, Puhl L, Gudat U, Woodworth JR, Weyer C, et al. Premixed formulations of insulin lispro. Activity profiles in type 1 diabetic patients. Diabetes Care 1999;22:865—6.
[16] Howey DC, Bowsher RR, Brunelle RL, Woodworth JR. [Lys(B28), Pro(B29)]-human insulin. A rapidly absorbed analogue of human insulin. Diabetes 1994;43:396—402.
[17] Livesey G, Wilson PD, Dainty JR, Brown JC, Faulks RM, Roe MA, et al. Simultaneous time-varying systemic appearance of oral and hepatic glucose in adults monitored with stable isotopes. Am J Physiol 1998;275:E717—728.
[18] Singhal P, Caumo A, Carey PE, Cobelli C, Taylor R. Regulation of endogenous glucose production after a mixed meal in type 2 diabetes. Am J Physiol 2002;283:E275—283.
[19] Mooy JM, Grootenhuis PA, deVries H, Kostense PJ, Poppsnijders C, Bouter LM, et al. Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: The Hoorn study. Diabetologia 1996;39:298—305.
[20] Wilinska ME, Schaller HC, Schaupp L, Pieber TR, Chassin LJ, Hovorka R. Hypoglycaemia reduces and insulin increases interstitial to plasma glucose ratio during physiological conditions in subjects with Type 1 diabetes. Diabetologia 2002;45(Suppl. 2):A310.
[21] Efron B, Tibshirani R. An introduction to the bootstrap. Chapman & Hall; 1993.
[22] Mari A, Oestergaard T, Nyholm B, Schmitz O, Ferrannini E. Model analysis of beta-cell function in first-degree relatives of type 2 diabetics and its relation to insulin resistance. Diabetes 2002;51:1473.
[23] Wilinska ME, Schaller HC, Schaupp L, Pieber TR, Hovorka R. Modelling interstitial glucose kinetics in subjects with type 1 diabetes during physiological conditions. In: Ghorbel FH, editor. Proceedings of the Second Joint EMBS-BMES ConferenceIEEE, Houston; 2002; p. 228—9.
[24] Toschi E, Camastra S, Sironi AM, Masoni A, Gastaldelli A, Mari A, et al. Effect of acute hyperglycemia on insulin secretion in humans. Diabetes 2002;51:S130—133.
[25] Camacho EF, Bordons C. Model predictive control. Springer-Verlag; 1999.
[26] Parker RS, Doyle III FJ, Peppas NA. A model-based algorithm for blood glucose control in type I diabetic patients. IEEE Trans Biomed Eng 1999;46:148—57.
[27] Cobelli C, Federspil G, Pacini G, Salvan A, Scandellari C. An integrated mathematical model of the dynamics of blood glucose and its hormonal control. Math Biosci 1982;58:27—60.
[28] Salzsieder E, Albrecht G, Fischer U, Freyse EJ. Kinetic modeling of the glucoregulatory system to improve insulin therapy. IEEE Trans Biomed Eng 1985;32:846—56.
[29] Hovorka R, Andreassen S, Benn JJ, Olessen KG, Carson ER. Causal probabilistic network modelling — An illustration of its role in the management of chronic diseases. IBM Syst J 1992;31:635—48.
[30] Carson ER, Cobelli C, Finkelstein L. The mathematical modeling of metabolic and endocrine systems. New York: Wiley; 1983.
[31] Schaupp L, Ellmerer M, Brunner GA, Wutte A, Sendlhofer G, Trajanoski Z, et al. Direct access to interstitial fluid in adipose tissue in humans by use of open-flow microperfusion. Am J Physiol 1999;39:E401—408.
[32] Goriya Y, Ueda N, Nao K, Yamasaki Y, Kawamori R, Shichiri M, et al. Fail-safe systems for the wearable artificial endocrine pancreas. Int J Artif Organs 1988;11: 482—6.
[33] DCCT Research Group. Epidemiology of severe hypoglycemia in the diabetes control and complications trial. Am J Med 1991;90:450—9.
[34] World Health Consultation, Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Part 1. Diagnosis and Classification of Diabetes Mellitus. World Health Organisation, Geneva, 1999.
[35] Service FJ, O'Brien PC. The relation of glycaemia to the risk of development and progression of retinopathy in the diabetic control and complications trial. Diabetologia 2001;44: 1215—20.
[36] Schaller H, Bodenlenz M, Schaupp L, Plank J, Wach P, Pieber TR, et al. MPC algorithm controls blood glucose in patients with Type 1 diabetes mellitus under fasting conditions using IV-SC route. In: Diabetes Technology Meeting. San Francisco; 2001; A48.