A randomised-controlled clinical trial methodology for evaluating the teaching utility of interactive educational diabetes simulators

P. Tatti* and E.D. Lehmann**

ABSTRACT. AIDA is an interactive educational diabetes simulator which has been made available without charge on the Internet. Since its launch on the World Wide Web in 1996 over 58,000 people have visited the AIDA Web site (http://www.2aida.org) and over 17,500 copies of the program have been downloaded from there free-of-charge. The AIDA software is believed to be of use in recreating clinical (diabetes) situations for interactive simulation. However, despite its widespread usage, its actual utility for supporting the education of patients with Type I diabetes mellitus remains to be objectively demonstrated in a randomised-controlled clinical trial setting. This paper describes a prospective, randomised-controlled trial (RCT) methodology for formally evaluating the educational utility of an interactive diabetes simulator, like AIDA. The protocol makes use of two study arms, each receiving different educational interventions. During lessons, Arm A of the study will be exposed to the AIDA simulator (the active intervention), while Arm B (the control group) will benefit from conventional educational methods using standard presentations with slides and transparencies. Six lessons will be held for each study arm (one per week). At the beginning and end of the study self-monitoring blood glucose (SMBG) data will be collected, details of any hypoglycaemic episodes recorded, and assessments made of HbA$_{1c}$. Participants will also be required to complete a detailed questionnaire to assess their self-confidence, quality of life and metabolic control, attitudes towards SMBG, and knowledge about insulin dosage calculation. Comparisons will be made between Arm A and Arm B using unpaired statistical analyses. A partial cross-over study design is also proposed whereby subsequently the control group will be exposed to the AIDA simulator during a further 6-week course of lessons. This will ensure that the maximum number of subjects will eventually receive the active intervention, and will also allow further within group paired analyses to be applied (with greater statistical power). An initial evaluation study using this RCT approach has just recently commenced in the Ospedale di Marino in Marino (Rome), Italy.


INTRODUCTION

Since the discovery of insulin there has been a continuous effort to improve the quality of life of patients with diabetes. As with most other chronic diseases, therapeutic education has become a mainstay of treatment. Initially education focused on the dietary and behavioural aspects of the disease. However with the advent of more intensive insulin regimens, which allowed more flexibility, the need for a more structured education program, including detailed information on insulin dosage adjustment, has become increasingly important. Another impetus comes from the Diabetes Control and Complications Trial (DCCT), which demonstrated that an intensi-
fied insulin regimen (1), including many education sessions, can improve the prognosis in Type 1 diabetes mellitus (T1DM). Unfortunately the emphasis on therapeutic education has often been academically oriented, rather than necessarily being tailored to the specific needs of patients with diabetes (2). Connected with this, it can sometimes be easier to teach about the disease process and the pathophysiology of diabetes than about the practical things (insulin-dosage adjustment, balancing insulin and diet, etc.) that patients really need to know.

Many authorities, including the World Health Organization (WHO) (3), have called attention to the need to focus education on the more basic needs of the diabetic population. The importance of therapeutic patient education at a time of escalating costs and shrinking resources has also been emphasised (4). However there is considerable confusion in the medical literature about the goals and the methods applied for therapeutic education (5). Furthermore, conventional education in the clinical setting can meet with many obstacles, such as lack of human resources, the tradition and culture of health care professionals, insufficient team work, conservatism, difficulty in assuring valid evaluation, and lack of educational and financial resources (3).

The recent introduction of computer simulations in therapeutic education may assist in this respect. Computer simulations have performed well in many fields, such as civil aviation, the military, engineering, finance, and physiology (6, 7), where they are used extensively, but their use in therapeutic patient education is relatively new and their advantages are still untested.

Furthermore, while there is increasing interest in the application of information technology in diabetes care (8, 9), attention seems to have focused to date rather more on the development of the software than on evaluation issues and actually demonstrating a clinical benefit from the use of such computer programs (10).

AIDA is an innovative educational diabetes simulator which can be used for demonstration and teaching purposes to simulate the effects of changes in insulin therapy and diet on the blood glucose (BG) profile of a ‘typical’ patient with T1DM (11, 12).

While other interactive simulators of glucose-insulin interaction in diabetes have been described in the literature (13-16), to date none of these have been distributed widely, nor extensively used. By contrast, the AIDA software can be downloaded free-of-charge from the Internet (http://www.2aida.org), as a non-commercial contribution to continuing diabetes education (17). Up to the end of October 2000 over 17,500 copies of the program were downloaded for free from this Web site.

Detailed descriptions of the AIDA program and how it works can be found elsewhere in the literature (12, 17-20). Briefly, the AIDA simulator incorporates a compartmental model which describes glucose-insulin interaction in patients completely lacking endogenous insulin secretion. It contains a single extracellular glucose compartment into which glucose enters via both intestinal absorption and hepatic glucose production. The AIDA model also contains separate compartments for plasma and “active” insulin (13), the latter being responsible for glycaemic control while insulin is removed from the former by hepatic degradation. Figure 1 summarises the anatomical basis of the compartmental model (19), the mathematics of which have been described elsewhere (18). Full details of the AIDA model are also accessible from within the AIDA software package, and can be viewed and printed separately from the Internet (http://www.2aida.org/technical).

EXAMPLE CASE

The AIDA software comes with 40 educational case scenarios as standard, each representing a ‘snapshot’ of the metabolic status of a typical patient with respect to T1DM. It is easy for users to add or create further case scenarios, as required.

Figure 2A shows a baseline simulation for one of these 40 example case scenarios. The only information that the user is provided about this patient is that: ‘This woman is on a four times daily insulin regimen, taking three ‘shots’ of a short-acting preparation before each of the main meals, with an intermediate-acting preparation before going to bed. She injects using an insulin pen – but has not yet managed to stabilise her glycaemic control. She tends to eat a lot more towards the end of the day and so finds herself going markedly hyperglycaemic overnight. How might you control these extremely raised blood sugar levels during the night, without sending her ‘hypo’ as a result?”

The upper graph shows the simulated BG data for this case, while the lower graph provides a composite display of information regarding insulin and
A methodology for evaluating educational diabetes simulators

**GLUCOSE MODEL**

**BRAIN**

**EXTRACELLULAR COMPARTMENT**

**PERIPHERY**

**LIVER**

**KIDNEY**

**URINE**

**PATIENT SPECIFIC MODEL PARAMETERS**
1. Body weight: 70 kg
2. Renal threshold of glucose (RTG): 9.0 mmol/l
3. Creatinine clearance rate (CCR): 100 ml/min
4. Hepatic insulin sensitivity ($S_h$): 0.5
5. Peripheral insulin sensitivity ($S_p$): 0.5

**INSULIN MODEL**

![Graphs showing insulin levels over time](image)

Fig. 1 – Anatomical basis and physiological functions of the AIDA model. AG: arterial plasma glucose; BG: blood glucose; CHO: carbohydrate; GAR: glucose absorption rate; IIGU: insulin-independent glucose utilisation; NHGB: net hepatic glucose balance; PGU: peripheral glucose utilisation; RBC: red blood cell; UGER: urinary glucose excretion rate. Modified from (19).
short- and intermediate-acting ones, or perhaps try the case scenario with a different ‘pen regimen’ taking a longer-acting insulin preparation at night. The list of possibilities is endless: a near infinite number of simulations can be performed with AIDA.

As can be seen in Figure 2A, the “virtual diabetic patient” in this example tends to have quite high BG levels overnight. Figure 2B shows a simulation of one possible way of reducing this hyperglycaemia, the effect of moving the intermediate-acting injection from bedtime (21:30 h) to before supper (18:40 h) and decreasing the bedtime snack from 40 g of carbohydrate to 20 g being shown. While these adjustments do not bring the overnight BG levels fully under control, they do substantially reduce the carbohydrate intake. The distribution of the meals eaten can also be seen in this panel along with the four times daily regular (short-acting) and intermediate-acting insulin regimen. Superimposed on these graphs are predicted steady state BG and plasma insulin profiles as calculated by the AIDA model.

Having performed a baseline simulation users can then alter any of the input variables to simulate the glycaemic effects of such changes. For example a user could simulate what would happen to a hypothetical patient’s BG profile if the bedtime intermediate-acting insulin dose was increased by 2 U, or the injection time moved earlier, or the bedtime snack shifted later, or the carbohydrate content of supper decreased by 10 g. A user could transfer the patient to a different preparation in place of the previous
night-time hyperglycaemia that was previously occurring. The AIDA software can simulate a wide variety of other insulin dosage and dietary adjustments. However, it should be stressed that the purpose of AIDA is to create a learning environment for communicating and training intuitive thinking when dealing with insulin dosage, dietary and lifestyle adjustments. The software is not meant for individual patient glycaemic prediction or therapy planning (12). In this respect AIDA appears most of use for recreating clinical situations rather than trying to predict the best outcome.

Further examples of the application of AIDA as an educational tool can be found elsewhere in the literature (12, 17, 20), and a full demonstration can be viewed on-line at, or downloaded without charge from, the AIDA Web site. User comments about the AIDA software (from a wide range of end users – eg patients, their relatives, health-care professionals and students) can also be found elsewhere in the literature (17, 21, 22), as well as on the Internet at http://www.2aida.org/reviews

PREVIOUS STUDIES

To date, demonstrating a benefit from the use of interactive educational computer-based diabetes simulators has proven difficult. One system, called DIABLOG, was assessed in a preliminary manner by 22 patients with T1DM, but only subjective impressions of the users were reported (23). In a preliminary evaluation of a diabetes computer game, called "Captain Novolin"™, which incorporated simulation functions, 23 children with diabetes aged 6-16 yr who had used the game and one of their parents were interviewed by the system developers (24). When given a choice of educational material about diabetes, 22 out of 23 children chose to have patient education from a computer game and only one child chose a video tape. After playing the game for half an hour, more than two-thirds of the children wanted to continue playing the game (24). However, once again, no objective data were available to show a benefit from the game.

Baldwin et al (25, 26) used questionnaires to assess the potential benefits of the Hauser interactive diabetes simulator (27). In this study patients were taught insulin-dosage adjustment by a diabetes educator using either the simulator, or a pen and paper. No significant differences in outcomes were observed between the two cohorts (25, 26). However, educators are very well trained, and in particular are trained to do their job without a computer. Therefore, perhaps adding a computer to their existing sessions may not make too much difference to measures of patient benefit of such interventions. However, it should be noted that these studies (25, 26) were only carried out for a relatively short period of time (1 month) in quite heterogeneous and small groups of patients (with both T1DM and Type 2 diabetes) of quite widely varying ages. Unfortunately the program’s textual interface (which might have been less intuitive for some patients), as well as the short timescale, and heterogeneous and small sample cohorts, may also have impacted on obtaining a positive outcome in these studies.

A preliminary attempt to evaluate the educational benefits of “Sarimnet” (15), an interactive graphical diabetes simulator, was undertaken in 1991 (although the results were only published in 1995) (28). The authors reported that only 11 out of 58 diabetic teenagers wanted to participate in the study. These diabetic teenagers were educated in 4 computer lessons and evaluated with respect to metabolic control, emotional adjustment, locus of control, self-esteem and ability to discuss treatment problems. It was not possible to recruit a control group. In a fairly small group there were significant improvements in locus of control, self-esteem, knowledge and diabetes-related stress. The computer training for the most part was considered to be enjoyable and of great value. Interestingly, the authors reported that a side-effect of using their program, for some patients, was an increased level of guilt and alienation from medical professionals. However of note, the participants generally regarded the computer simulator as a source of inspiration rather than as an instrument for calculating the optimal insulin regimen (28). The authors concluded that although young people get more and more used to computers, still only a minority are attracted to this type of education. It should be stressed, however, that this study was undertaken in 1991. Since then a massive expansion in the home and work use of computers has taken place. Furthermore, this study, like the studies referred to above (23-26, 28), suffered from very small numbers (n=11) which precluded the identification of any statistically significant effects (20).
One of the more encouraging evaluations of an educational computer program incorporating simulation functions in diabetes care has involved "Packy & Marlon"™ (Raya Systems Inc., California, U.S.A.), a role-playing Super Nintendo video game in which children manage the diet and insulin of two elephants who have diabetes (29). To optimise the educational benefits of this game players can select insulin plans that match their own. The scenario is a diabetes summer camp which has been raided by rats. The two elephants, Packy and Marlon, need to defend themselves by blasting the attacking rodents with peanuts and water from their trunks. They also need to find food and supplies, remembering to eat healthily, regularly check their BG levels, and take their insulin (29).

"Packy & Marlon"™ was assessed in two centres in a 6-month randomised-controlled trial, in a cohort of 59 children with T1DM. Half the cohort received the diabetes game to use at home as much as they liked, while the other half (the control group) received a video game with no health-care content. While significant improvements in HbA1c were not demonstrated, the authors highlighted that the patients were reasonably controlled at the start (mean baseline HbA1c: 8.3-8.5%) and therefore the study quite possibly ran into a "ceiling effect" (10). Clearly to overcome this problem further randomised-controlled trials with diabetic children with more usual (poorer) glycaemic control would be required. Notwithstanding this, benefits were reported in diabetes self-efficacy, communication with parents about diabetes, and self-care behaviour in the children who received "Packy & Marlon"™. There was also a decrease in unscheduled urgent doctor visits, a finding which, if confirmed by larger studies, would be most encouraging for children with diabetes. Furthermore the more a child played the diabetes game, the more friends they spoke with about diabetes (30).

The aforementioned studies, to our best knowledge, are the only published evaluation reports to date of diabetes simulation software for educational usage. Part of the reason for this relative paucity of published data undoubtedly is the work required to actually document a benefit from the use of such programs. However, another reason is likely to be the lack of a standardised protocol for evaluating such simulation tools, as well as uncertainty about the best methods to actually use for the evaluation (17).

**CURRENT STUDY**

The purpose of this report is to document, in detail, a randomised-controlled clinical trial protocol which can be widely applied for the evaluation of such interactive educational diabetes simulation software. We are aware of previous experience with more standard (non-computer-based) educational interventions that have demonstrated an improvement in patient knowledge without an improvement in metabolic control (2). Therefore it is important to highlight in advance that our study design does not seek to identify an increase in theoretical (book-type) knowledge among participants. Rather the aim is actually to establish how practical benefits in terms of increased glycaemic control and patient self-confidence can be accomplished using computer-based diabetes simulations. AIDA serves well as a test platform for this approach because of its widespread availability via the Internet. Also it is comprehensiveness, interactive graphics, and the possibility to simulate many different case scenarios – with direct involvement of patients with diabetes – makes it a potentially useful teaching aid. Furthermore, the 40 standard case scenarios included within the AIDA database can be used to illustrate a selection of different educational points, as will be highlighted below.

Therefore the AIDA simulator has been chosen as the primary software for evaluation. However, it is hoped that other simulation programs could also be assessed equally well using this approach. Furthermore, by documenting the protocol in this way, perhaps more objective comparisons will become possible between different diabetes simulators.

An initial evaluation using this study design has just recently commenced in the Diabetes Clinic at the Ospedale di Marino, in Marino (Rome), Italy.

**PROTOCOL**

The outcome measures which this protocol is intended to evaluate are set out in Table 1. To demonstrate an improvement in these we have designed a prospective, randomised-controlled trial methodology consisting of conventional education sessions vs computer-simulation sessions, with a partial cross-over design to maximise the numbers of subjects exposed to the simulator. Using such an approach, both between-group and within-group comparisons and statistical analyses will be possible. Figures 3 and 4 summarise the overall study design.
Table Main (and secondary) outcome measures of study.

Main outcome measures

1. Improve knowledge of how to tailor the insulin dose to the specific needs of the subject
2. Improve morning and post-prandial self-monitoring blood glucose (SMBG) data
3. Improve forward thinking (ability to answer "what-if" type questions)
4. Improve well being and self-confidence
5. Reduce the number of hypoglycaemic episodes
6. Improve HbA1c
7. Achieve subject feeling of "empowerment"

Secondary outcome measures

8. Increase social behaviour and interpersonal co-operation about diabetes
9. Help to increase confidence with a computer
10. Increase physiological knowledge
11. Establish sample sizes required for a possible future larger-scale, multi-centre study

SUBJECTS

For the initial (pilot) evaluation study we are planning to enrol 24 adult subjects with T1DM, and randomly assign them to one of two study arms [Arm A: simulator; Arm B: conventional education (control group)]. The subjects will be recruited from the Diabetes Clinic at the Ospedale di Marino, Marino (Rome), Italy.

The self-reported incidence of symptomatic hypoglycaemic episodes in this clinic is about 10-20 episodes per month (in 65 subjects). If we take an arbitrary cut-off of 70 mg/dl (3.9 mmol/l) in the self-monitoring BG (SMBG) data, as representing a low BG level, then the number of low BG readings rises to nearly 100-150 per month. It is important to note that almost all the patients with T1DM intend-
ed to participate in the initial study will have an HbA1c <8.5%, as they all undergo intensive therapy (4 injections per day, regular SMBG, with close follow-up). This tight glycaemic control can lead to an unwanted number of hypoglycaemic episodes. Therefore, identifying whether education with the simulator can reduce the number of unwanted "hypos" is an important goal in this intensively treated cohort.

Inclusion and exclusion criteria for the study are summarised in Table 2. Every subject will be required to give their written informed consent before

---

Table 2 - Study inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Type 1 diabetes mellitus with &gt;6-year duration</td>
</tr>
<tr>
<td>2. Use of at least 0.7 U of insulin per kg of body weight</td>
</tr>
<tr>
<td>3. Body mass index ≤26 kg/m²</td>
</tr>
<tr>
<td>4. 18 years ≤ age ≤ 50 years</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria*</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Presence of serious diabetic complications or other disease that impairs visual capacity or ability to participate in the study</td>
</tr>
<tr>
<td>b. Lack of dependability or a high level of probability that the subject will not attend all the lessons</td>
</tr>
<tr>
<td>c. Pregnancy</td>
</tr>
<tr>
<td>d. Unwillingness to participate</td>
</tr>
<tr>
<td>e. Any history of cheating with self-monitoring blood glucose</td>
</tr>
<tr>
<td>f. Prior experience with a diabetes simulator</td>
</tr>
</tbody>
</table>

*Items 1-3 should be sufficient to exclude consistent residual insulin secretion (as in Type 2 diabetes mellitus).

*subjects will also be excluded from the study (and all subsequent analyses) if, once recruited, they miss more than two lessons.
entering the study. Local Ethical Committee approval for the study has been obtained.

**PRE-RANDOMISATION (BASELINE) PHASE OF STUDY**

The pre-randomisation (baseline) phase of the study is summarised in Figure 3. Upon recruitment every subject will be asked to keep a log of any symptomatic hypoglycaemic episodes, and collect their SMBG data for one week before breakfast (fasting) and 2 hours after lunch and dinner. Subjects will only be included in the study if they have collected at least 75% of the SMBG data requested. The patients’ logbook entries will be cross-verified with the memory from their reflectance meter. Since an obvious question can be raised about the level of knowledge and experience of the study subjects, it will be necessary to collect certain baseline demographic data about each participant from the hospital/clinic information system. These data are summarised in Table 3.

After this, the subjects will be randomly assigned to either Arm A or Arm B of the study. The assignment will use a process of stratified randomisation that takes into account hypoglycaemia/low BG risk and HbA1c level. The purpose of the stratification is to ensure that Arm A and Arm B of the study have similar numbers of subjects in relation to key variables which might influence the outcome.

Subjects will be stratified into 4 groups according to: 1) low HbA1c level, low “hypo” risk; 2) high HbA1c level, low “hypo” risk; 3) low HbA1c level, high “hypo” risk; and 4) high HbA1c level, high “hypo” risk (Fig. 3).

---

**Table 3 - Baseline demographic data to be collected about each subject from the hospital/clinic information system.**

- Age
- Body mass index
- History of hospitalisation for diabetic ketoacidosis
- Duration of diabetes
- Previous experience with diabetes education sessions
- Degree of metabolic control*
- History of hypoglycaemic episodes
- Job and educational level

*Average of previous 12 months’ HbA1c measurements (also used for stratified randomisation)

For the purpose of stratification each subject’s individual HbA1c level will be assessed as the average (arithmetic mean) of their previous 12 months’ HbA1c data (Table 3). A low HbA1c will be defined as less than the median (50th percentile) HbA1c level for the whole cohort, while a high HbA1c will be defined as greater than or equal to the median HbA1c level of the cohort.

“Hypo” risk will be categorised based on the one-week SMBG data collection. High “hypo” risk will be defined as any symptomatic hypoglycaemic event(s) or recorded SMBG data <70 mg/dl (3.9 mmol/l) during that week. Low “hypo” risk will be defined as no symptomatic hypoglycaemic episodes and recorded SMBG data >70 mg/dl (3.9 mmol/l) during that week.

**Method of randomisation**

In order to ensure complete separation of subject recruitment and study execution from the randomisation process, an external randomisation service has been set up via the AIDA Web site (http://www.aida.org/random). This allows subject details (patient name, hospital/study number, gender, age, number of symptomatic “hypo”s and average HbA1c level) to be entered for each study participant. It is intended that entries should be made by the physician or nurse who recruited the subjects, categorised according to low/high “hypo” risk and low/high HbA1c levels, as outlined above. Data entered in this way can then be automatically submitted across the Internet, via e-mail, to a co-investigator (E.D.L.) in London, England. Using dedicated software written specifically for this purpose (incorporating a random number generator) it will be possible for this co-investigator to independently randomise participants either to the AIDA or control groups, randomisation codes being returned to the referring physician/nurse by e-mail.

In this way there should be a complete division between the recruitment, enrolment and teaching of subjects, and their initial randomisation to either Arm A or Arm B of the study.

**MAIN PHASE OF STUDY**

The main phase of the study is summarised in Figure 4. Just before entering this phase the enrolled subjects will have their baseline HbA1c measured, and they will fill in a study questionnaire.
Study questionnaire

The main study questionnaire is designed to test the subjects’ competence with insulin dosage calculations and their level of well-being and self-confidence rated on a scale (see Appendix). The questionnaire is divided into 7 sections covering baseline questions, plus questions about: 1) patient self-confidence; 2) quality of life and metabolic control; 3) social and emotional impact of diabetes on lifestyle; 4) attitudes towards SMBG; 5) prior knowledge about diabetes; and 6) "what-if" type questions and knowledge about insulin-dosage adjustment. The first sections of the questionnaire are from validated sources, section 1) using a modified Grossman self-efficacy scale; and sections 2-5) from the Diabetes Care Profile, from the Michigan Diabetes Research and Training Center (31) (see Appendix). The baseline questions and section 7) have been written specially by the study authors. Furthermore the general selection of the individual questions reflects the experience of the authors. For example, since urine testing is no more in great use, all questions concerning glycosuria have been omitted. However, it will be self-evident that the selection of such questions should reflect the local experience of the area where the study is conducted. Therefore the questions chosen for this study are based on the particular experience of the protocol authors regarding the most common pitfalls in diabetes self-management.

The questions are closed-ended and the options presented come specifically from a small pre-test pilot study involving T1DM and insulin-treated Type 2 patients with diabetes. When it was felt that the options were not exhaustive a free text line was included. However, most questions use a summative (Likert) scale. Each questionnaire will be linked to an individual participant by a study number, and not by the name of the subject. The correspondence of name to study number will be registered separately by a nurse and will remain blinded till the end of the study.

However, before any studies using this protocol actually commence at least two experienced diabetologists should agree on the best local answers – and pre-rank the answers on a scale, from totally inappropriate, to reasonable, to a good option, to the best option – or any scale like this.

In addition to being used to help assess the outcome, it is intended that data from the baseline questionnaire should also be used to assess any significant difference(s) between the two study arms (A and B) in terms of key variables, ie whether stratified randomisation has worked. This will effectively permit a pre-study test to confirm that the two arms of the study are not significantly different in any key respects, and particularly that they are well-randomised in terms of HbA1c levels and "hypo"/low BG risk.

Interventions

The two arms of the study will receive different interventions. Arm A will use the AIDA simulator as described below, while Arm B will benefit from conventional (standard) educational methods (verbal communication with slides and transparencies).

The aims of the lessons provided to the two study arms will be identical, as outlined in Table 4. Where there is more than one teacher, the expert(s) who will teach the subjects will have a session of nearly 2 hr with each other, before starting the study, to simulate the teaching process and to homogenise their strategy and behaviour.

The study subjects will be randomised into groups of 6 subjects each (with multiple groups planned). All the study subjects will be asked not to use the AIDA simulator on their own to avoid the possibility that one “over-performer” can lead the group, or that the results can be biased in any way, for instance by someone in the control group (Arm B) accessing the simulator at home.

Before the start of the main phase of the study, all subjects will be asked to bring with them to lessons their own SMBG data, insulin regimen, and carbohydrate intake details for possible use during classes. At the end of each lesson, subjects will be given a short, separate generic one-page questionnaire to complete regarding their thoughts on the lesson,
and whether they felt that they benefited from it. All subjects will go through 6 lessons, one each week. In the description of the two study arms which follows, more details are given about the simulator arm (Arm A) simply because this is the novel intervention being tested. However, it is intended that the duration of the lessons will be the same for both study arms (approximately 65-70 min).

ARM A (SIMULATOR ARM) – PHASE 1
The participants will not interact directly with the computer, rather the teaching expert(s) will serve as "wizards" undertaking all interactions with the computer. This will circumvent any problems if the participants are not computer-literate or confident to use a computer themselves. Also this approach should help prevent any "learning curve" effect where it takes the study participants a number of lessons to become fully familiar with the program's functions. Furthermore, using a "wizard" approach should prevent any language difficulties as the software is written in English, while the participants' mother-tongue – certainly for the initial evaluation study – will be Italian. In this respect, where required, the "wizards" will also serve as translators. For all simulation lessons two computer screens will be made available, linked to the same PC, ensuring that all study participants will have a good view of the computer displays. At the end of each lesson the number of simulations run during that lesson will be recorded. The topics to be covered during each lesson are summarised in Table 4.

**Lesson I:** This lesson will be spent familiarising the subjects with the aims of the simulator, to explain the principles of operation, and some theoretical details of the "what-if" type thinking in diabetes care (10 min). The limitations of the simulator will also be put forward clearly to the subjects; 15 min will be spent explaining the kinetics of the insulin used. Particular emphasis will be placed on the difference between the analogue (rapidly-acting) insulin (eg lispro) and short-acting, regular insulin. Then the group will go through the first simulation. Case scenario number 0030 from the AIDA database will be used for this (Fig. 2). The remaining 15 min of the lesson will be spent on the interactive simulation of a case presented by one of the study participants. While it may seem early to involve the subjects to such an extent, from the first lesson, in the authors' experience, the participants will be eager to present their problems, which can be a great stimulus for them to come to the lessons. Indeed, the main reason they are likely to attend will be to solve their own particular problem(s). It is quite possible that actually taking part in the study will only be a secondary reason.

After the presentation of a case, a discussion will follow and each participant will present his/her suggestion(s) as to possible solutions. Attempts will be made to focus the discussion on the rationale behind such suggestions.

Before leaving the class each participant will be asked to prepare his/her own case scenario to be used during one of the following lessons. This is intended to involve the subjects more in the process and to save time during future lessons.

The general scheme of lessons II to VI will involve a 10-min introduction, followed by a 20-min simulation of the pre-selected case (*vide infra*), and then a further 20 min to interactively simulate a case presented by one of the study participants. Therefore, during each of these lessons at least two separate cases will be presented. As before, after the presentation of a case a discussion will follow and each participant will present his/her suggestion(s) as to possible solutions. The discussion will be focused on the rationale behind such suggestions.

**Lessons II and III:** The group will go through 2 other scenarios from the AIDA database (cases 0010 and 0027). Throughout it will be made clear that the computer simulation is not offering a definitive solution, but rather just one of the options to be discussed. At least 20 min of the second and 10 min of the third lesson will be spent discussing the effect of physical activity, even though the simulator at present does not cater for this.

**Lesson IV:** This lesson will deal with hypoglycaemia. The situations predisposing to "hypos" and interventions with oral carbohydrates will be simulated in detail. During this lesson case scenario 0035 from the AIDA database will be presented.

**Lesson V:** This lesson will focus on shift work and travel abroad. Case 0037 from the AIDA database will be used for simulation. As before, a further case for simulation will come from the experience of one of the study participants.
Lesson VI: This lesson will focus on unexpected events or missed insulin injections. Case scenario 0034 from the AIDA database will be used for simulation.

Ten min will be allowed at the end of each lesson to complete a short, generic one-page questionnaire (see Appendix).

ARM B (CONTROL ARM – CONVENTIONAL EDUCATION) – PHASE 1

For this group the content of a standard textbook will be used as a reference (32). The principles are the same as described above for Arm A and the topics will be the same as set out in Table 4, but this group will not use the AIDA simulator.

The general scheme of these lessons will involve conventional (standard) education sessions on a given topic, using slides and transparencies, followed by discussion, questions and then a recap.

Lesson I: The first 10 min will be spent providing an introduction to the 6-lesson course. The next 15 min will be used to explain the kinetics of the different insulin formulations. The remaining time will be spent illustrating how to adjust the insulin dose.

Lesson II: This lesson will be spent discussing the use of SMBG and the role of the liver in the fasting state.

Lesson III: This lesson will discuss the role of the renal threshold of glucose as well as introduce the topic of physical activity.

Lesson IV: This lesson will cover hypoglycaemia, its causes, and how to correct it.

Lesson V: This lesson will be devoted to the timing of meals and apportionment of carbohydrates, as well as shift work and overseas travel.

Lesson VI: The final lesson will cover special occasions/unforeseen events, eg missed insulin injections.

As for Arm A, at the end of each lesson the participants will be given a short, generic one-page questionnaire to complete regarding their opinion on the lesson.

CROSS-OVER PHASE – PHASE 2

At the end of the control arm (Arm B – Phase 1) there will be a gap, once Phase 1 is complete, of 4 weeks, following which the subjects will fill in the study questionnaire, have their HbA₁c re-measured, and then enter the cross-over phase of the study. This will entail a further cycle of 6 lessons, exactly the same as Arm A (with the simulator).

The purpose of this partial cross-over phase is to expose the control (standard education) group to the simulator, and therefore increase the final sample size actually receiving the active intervention. In addition to allowing between-group analyses, this approach will also allow further comparisons to be made within groups of the same subjects, and therefore allow more paired statistical analyses to be used, with their associated greater statistical power.

FINAL OUTCOME MEASURES

At the end of the study, after the interventions (both of Phase 1 and Phase 2), the participants will be required to:

1) collect SMBG data for 1 week, before breakfast (fasting), and 2 hr after lunch and dinner;
2) have their HbA₁c measured (one month after the end of the intervention);
3) provide a record of any symptomatic hypoglycaemic episodes during the study;
4) complete the study questionnaire to assess their self-confidence and well-being (rated on a scale), together with their knowledge of “what-if” type problem solving in diabetes care. An additional page of feedback comments will be sought from the participants in the simulator arm (Arm A) concerning what they thought about the AIDA simulation software (see Appendix).

The recording of symptomatic hypoglycaemic episodes will be differentiated from biochemical “hypos” [recorded low BG levels <70 mg/dl (3.9 mmol/l)].

STATISTICAL ANALYSES

The evaluation of the questionnaire data will be made by a person blind to the identity of the proband, the arm of the study, and phase of the trial. This point may be relevant especially for the questions with open-ended answers.

Assuming that data are distributed normally, the sta-
tistical tests that appear most useful are paired sample t-tests for analyses between Phase 1 and Phase 2 for the same subjects, and independent sample t-tests for analyses between Arm A and Arm B for different subjects. However, within-group comparisons (paired statistical analyses) clearly could be used for most analyses, as each subject will have their own baseline (control) data. As such, these tests are likely to be the most powerful and useful until we have a sufficiently large sample size and a priori power calculations to support the identification of a given sized effect. Wilcoxon signed-rank tests and Mann-Whitney U-tests will be applied for non-normally distributed data. A general linear multivariate regression model may also be used. Furthermore Likert scale data can be readily evaluated with the Cronbach α test (33).

Subjects who miss more than 2 lessons will automatically be excluded from the study and will not be included in any subsequent statistical analyses.

STUDY LOGISTICS

It is well recognised in education that the smaller the class size the greater the educational opportunity that can be offered to participants. For this study class sizes also need to be kept small in order to allow each subject to present their own data during one of the 6 simulation lessons. For this reason the class size will be limited to 6 subjects per class for both Arm A and Arm B. It is intended for subjects to have one lesson per week, thus it will take 6 weeks to run a complete course of 6 lessons.

Due to local time constraints and the availability of teachers, it may be necessary to run only one lesson for Arm A – and one lesson for Arm B – each week. This means that more than one 6-week course of lessons will be required to get all the recruited subjects through the lessons.

DISCUSSION

This report describes a prospective, randomised-controlled clinical trial protocol for evaluating the teaching utility of interactive educational diabetes simulators. In designing this approach it has been necessary to balance what is theoretically desirable with what is achievable in practice. Therefore, the practicalities of running such a study in a busy clinic teaching setting have had to be considered, as well as issues of balancing demands on time, and of teachers, and the number of questionnaires that participants can be expected to answer. Furthermore, in designing the current protocol we have also aimed to comply with the CONSORT randomised-controlled trial guidelines (34) so that any reports of the results of evaluation studies using this approach should more readily be CONSORT-compliant.

An important issue to be addressed by the initial evaluation study is that of the sample size. As this is the first time that such a study has been run, no one can be quite sure what the actual effect size will be. Therefore, determining the statistical power required for further studies will be an important outcome measure of the initial evaluation study using this approach.

One particularly novel aspect of the current protocol is the use of a “wizard” to interact with the computer. It is well recognised that there is often a learning curve to the use of computer software; and in this respect AIDA is no different from other computer programs. While after a small amount of use AIDA becomes actually remarkably easy to use, it is accepted that for a completely new user there may be an initial “learning curve” to become familiar with the user interface and all the program functions. Using a “wizard” to interact with the computer and the software circumvents many of these possible problems. Furthermore, any local resource difficulties about needing a large number of computers available in the diabetes clinic setting – for patients to access themselves – are circumvented by use of this “wizard” approach.

As computers become more powerful – and as technology and graphical user interfaces improve – interactive educational diabetes simulators are likely to become increasingly sophisticated (35). However, there remains a need to formally establish whether such simulators can actually contribute to diabetes education, and if so, how they may be best applied. The protocol described in this report sets out to establish this prospectively for the AIDA interactive educational diabetes simulator.

A further issue that was actively debated was how to arrange the pre-study stratified randomisation of participants into high and low risk groups. Initially it had been considered to use the incidence of “hypo”/low BG risk and educational status (level/class reached at school or university). However, it became apparent that the latter variable (educational status) would not be very discriminatory and therefore would not really help categorise or stratify the cohort. In this respect
it became clear that the HbA\textsubscript{1c} level would be a much more useful variable for stratification. Consequently, as outlined above, we have chosen to use historical HbA\textsubscript{1c} data (based on the average value from the preceding 12 months) for stratification as this will then permit the baseline HbA\textsubscript{1c} values, collected specifically for this study, to be used also to establish how well the stratification has actually worked (in what can be considered a pre-study test). An HbA\textsubscript{1c} measurement is generally regarded as providing an integrated value for the BG level over the previous 2-3 months (8-12 weeks). By contrast the interventions in this study – in both the simulator and control groups – only last 6 weeks. Therefore it is possible that any effects of the interventions on the participants’ HbA\textsubscript{1c} levels may only manifest themselves a while after the completion of the interventions (lessons). For this reason a decision has been reached to sample the participants’ HbA\textsubscript{1c} levels approximately 4 weeks after the end of the interventions (in both Phase 1 and Phase 2 of the study). This means that the follow-up HbA\textsubscript{1c} will be measured approximately 10 weeks (2.5 months) after the start of the lessons. Furthermore, the participants may take a while to put into practice the things that they have learnt during their lessons. Therefore, once again, a gap between the end of the lessons and the follow-up measurement of HbA\textsubscript{1c} could be helpful.

Connected with this, a 4-week “wash-out period” is also planned between the end of Phase 1 of the study and the beginning of Phase 2 (the cross-over phase). This is intended to reduce any “carry over” effect on BG levels for the control group (Arm B) – from the conventional lessons (Phase 1) to the simulator-based lessons (Phase 2).

A key requirement of a randomised-controlled trial protocol such as that described above is to avoid the introduction of any bias into the study. Clearly any differences between the experience of the simulator and control groups could potentially cause bias which might lead to differences in the study outcomes, which might actually not be attributable to the use of the simulator. For this reason it is important to minimise the introduction of bias and recognise where potential biases may arise. Given this, a number of aspects of the study design were actively debated – and considered at some length – before arriving at the final protocol described above. It may aid consideration of possible different methodological approaches to document the reasoning behind some of the decisions reached.

**Potential biases**

It was considered that where more than one expert (“wizard”) is involved, the two experts should switch between the two arms of the study so that the participants of both arms would get equal exposure to each expert – and so that there could not be any “bias” introduced by one expert teaching a particular group more (or better) than another. This might intuitively appear the most methodologically “pure” approach. However, it became apparent that the use of the simulator during a lesson might not only have a training effect on the subjects (as we hope to be able to demonstrate), but might also have a training effect on the teacher. Consequently, it might be possible for the teacher to end up modifying his/her own teaching as a result of using the simulator. This could incorrectly favour the conventional treatment arm who could actively receive benefits from the simulator through the teacher. Therefore, to avoid any such possible bias in the end we decided to have no switch of teachers between the two study arms.

Similarly, it will be important to ensure that subjects are randomly assigned to each course of lessons – ie there should not be significant differences or biases between the first time the lessons are run (the first 6-week block) and the subsequent runs (of the second/third 6-week blocks). To reduce the likelihood of this problem occurring the randomisation has been planned to take place at the beginning of the study – so there should not be any significant differences between characteristics of the study participants in the various 6-week blocks (or courses). It is also possible that with some “learning effect” for the teachers (“wizards”) their teaching may improve during the course of the study, with each teaching session where they make use of the simulator.

The only way of overcoming this potential problem is for the teacher using the simulator to be sufficiently familiar with the software from the outset of the study. Related to this we at least need to be aware of this possible confounder so that appropriate statistical tests can be applied to look for any relationship between possible outcome and the time that the course was run.

The need for generating a modified study questionnaire – different from the baseline version – had been considered in order to prevent any participant “learning effect” in actually filling in the questionnaire. However, it was necessary to balance any potential advantages of such an approach with the difficulties.
inherent in comparing and analysing the responses if the questions were different between baseline and follow-up. In particular it was debated how sure we could be that any changes identified in the questionnaire score would be solely due to increased participant learning – and not simply to the questions being different. Upon reflection we took the view that any “learning effect” in answering the questions was likely to be minimal after a 6-week study period. This is especially the case with regard to the participants’ knowledge and the “what-if” type questions as none of the lessons is actually directed to any of the particular questions posed in the study questionnaire. Furthermore, the majority of the questionnaire actually deals with issues of self-confidence and patient attitudes (see Appendix), which are unlikely to be directly influenced by having filled in the questionnaire before.

CONCLUSIONS
To evaluate the impact of education sessions with an interactive diabetes simulator on the well-being and short-term/intermediate metabolic control of T1DM patients, we have designed a study protocol using a rigorous and transparent methodology. The approach described above combines a prospective, randomised-controlled trial with a partial crossover design to maximise the number of subjects exposed to the simulator. The biggest limitation of this approach is likely to be the availability of teachers and the time required to actually run the lessons. Subjects will also need to be motivated as they will need to commit to 6-12 weeks of lessons. However, they do benefit from useful education sessions with conventional lessons and/or a simulator, so that should hopefully be enough to motivate them.

It should be evident that interactive simulations such as those highlighted here only provide one aspect of the education required by patients with diabetes. It is quite possible to envisage a study which incorporates components of the randomised-controlled trial described above, as part of the assessment of more standard educational diabetes interventions. In this respect it will be important not to restrict this protocol to use just with a teacher and 1-hr clinic lessons.

For instance, it is well recognised that there are not enough resources (trained diabetes educators, money, facilities, etc.) available for patients to be seen in clinic as often as might be regarded as ideal. For example, in the DCCT (1) patients in the intervention group were seen fortnightly and often contacted by telephone weekly (36). One issue appears to be whether information technology tools might be able to assist in bridging part of the gap between what was available in the DCCT v5 what is usually available routinely in day-to-day clinical practice.

Therefore other questions that we might wish to address in future evaluation studies include: does allowing a patient access to an interactive educational simulation program in between formal visits to a diabetes educator or clinic improve knowledge, confidence, self-management skills, or glycaemic control? To be able to judge whether such software can help in this role requires a rigorous evaluation methodology. Initially we are trialling the approach described above in a hospital clinic setting. However, in the longer term it is not difficult to envisage that the basic methodology of the questionnaire approach – combined where necessary with appropriate recording of clinical (biochemical) data – could be equally applicable for patients with diabetes using the simulator on their own at home (37), between clinic visits, or even possibly for health-care student evaluation (eg with medical, nursing, pharmacy, or diabetes educator students), who make use of the simulations for their own education.

SYSTEM AVAILABILITY
The AIDA diabetes simulation software continues to be developed and upgraded. The latest, new release of AIDA (v4.3) can be downloaded, completely free-of-charge, from http://www.2aida.org on the World Wide Web. The program runs on IBM PC or compatible 386/486/Pentium-based machines and requires approximately 2.2 Mb of hard disk storage space. AIDA can also be used on Apple Macintosh computers running PC emulation software such as Virtual PC or SoftWindows. People who wish to be automatically informed by email about updates and enhancements to the AIDA diabetes software range can subscribe (for free) to the AIDA registration/announcement list by sending a blank email note to: subscribe@2aida.org

Any readers who might be interested in collaborating by applying this protocol themselves in an evaluation of AIDA in their own unit(s), or who might be willing to refer patients for education sessions with the diabetes simulator, are invited to contact one of the authors.
REFERENCES


A methodology for evaluating educational diabetes simulators


APPENDIX

Three questionnaires have been designed for this randomised-controlled study. Limited space precludes publication of these here. However all three questionnaires can be freely downloaded and printed from the Internet from http://www.2aida.org/appendix and all are also available on request by e-mail from the authors (by writing to: www@2aida.org).

As described above the main study questionnaire is divided into 7 sections covering baseline questions, plus questions about: 1) patient self-confidence; 2) quality of life and metabolic control; 3) social and emotional impact of diabetes on lifestyle; 4) attitudes towards SMBG; 5) prior knowledge about diabetes; and 6) “what-if” type questions and knowledge about insulin dosage calculation.

A second generic one-page questionnaire has been developed for use at the end of each lesson. This is intended to establish if the participants felt more confident with the information provided during the lessons, and to get them to rank the information provided on a Likert scale from “not at all useful” through to “very useful”. Participants are also asked what areas they felt were not adequately covered by the lessons, allowing them to offer feedback.

The third questionnaire is intended for completion by participants at the end of Arm A of the study, after they have been exposed to the AIDA simulator. This contains 11 questions to establish what the participants thought of the simulator, and to identify ways in which they see the simulations being applied more widely in routine practice.