Clinical Requirements for Closed-Loop Control Systems

William L. Clarke, M.D.,¹ and Eric Renard, M.D., Ph.D.²

Abstract

Closed-loop (CL) therapy systems should be safe, efficacious, and easily manageable for type 1 diabetes mellitus patient use. For the first two clinical requirements, noninferiority and superiority criteria must be determined based on current conventional and intensive therapy outcomes. Current frequencies of hypoglycemia and diabetic ketoacidosis are reviewed and safety expectations for CL therapy systems are proposed. Glycosylated hemoglobin levels lower than current American Diabetes Association recommendations for different age groups are proposed as superiority criteria. Measures of glycemic variability are described and the recording of blood glucose levels as percentages within, above, and below a target range are suggested as reasonable alternatives to sophisticated statistical analyses. It is also suggested that Diabetes Quality of Life and Fear of Hypoglycemia surveys should be used to track psychobehavioral outcomes.

Manageability requirements for safe and effective clinical management of CL systems are worth being underscored. The weakest part of the infusion system remains the catheter, which is exposed to variable and under-delivery incidents. Detection methods are needed to warn both the system and the patient about altered insulin delivery, including internal pressure and flow alarms. Glucose monitor sensor accuracy is another requirement; it includes the definition of conditions that lead to capillary glucose measurement, eventually followed by sensor recalibration or replacement. The crucial clinical requirement will be a thorough definition of the situations when the patient needs to move from CL to manual management of insulin delivery, or inversely can switch back to CL after a requested interruption. Instructions about these actions will constitute a major part of the education process of the patients before using CL systems and contribute to the manageability of these systems.

J Diabetes Sci Technol 2012;6(2):444-452
**Introduction**

The development of practical and reasonably clinically accurate continuous glucose monitors (CGMs) has enabled the revival of the long-held dream of developing CL glucose control systems for the management of glycemia in persons with insulin-requiring diabetes.\(^1,2\) Reports of partial CL systems designed to halt infusion of insulin and avert predicted hypoglycemia; bi-hormonal systems that infuse insulin and/or glucagon to improve glycemic excursions; and CL systems with sophisticated model predictive control algorithms that have been shown to normalize postprandial glycemia, maintain euglycemia overnight, and prevent exercise-induced acute and late-occurring hypoglycemia, illustrate how rapidly and extensively research in CL technology has developed and expanded in less than 10 years.\(^3-9\) It is expected that one or more of these systems will emerge as a treatment option to conventional modes of treatment that neither normalize glycemia nor prevent many of the acute and/or long-term complications of diabetes. This article reviews some of the clinical requirements for CL systems and suggests possible ways in which these new systems might be evaluated.

**Efficacy and Safety**

A CL system, like any new treatment modality, be it a technical device, a pharmaceutical agent, or a revised surgical approach, must be both clinically efficacious as well as safe for patient use. At times, it may be difficult to separate these two requirements. For instance, in diabetes, in order for a treatment modality to be efficacious, it should have significant glucose lowering potential and simultaneously preventing an increase in or reducing the occurrence of low blood glucose (BG) levels. In order for a CL system to be safe for patient use, it should not be associated with extremely high BG levels and possible diabetic ketoacidosis (DKA), nor should its use incur an increased risk for severe hypoglycemia (SH; defined as loss of consciousness or inability to treat oneself). Thus, efficacy and safety are both part of the same clinical requirement, i.e., reducing hyperglycemic excursions while reducing the risk of SH. Efficacy should be categorized as superior or noninferior to current treatment outcomes. While it may be tempting to propose that any new product should be superior to currently available options in order to justify approval and use, superiority is not a Food and Drug Administration (FDA) requirement for approval for clinical use. Such a requirement could potentially prevent the timely availability of a treatment option that could have significant quality of life benefits to large numbers of people.

When evaluating the clinical efficacy and safety of a new treatment modality, it is important to determine the treatment modality with which it should be compared. It would be easy to compare CL control systems to conventional care as described in the Diabetes Control and Complications Trial (DCCT).\(^10\) Such a comparison would permit assessment of both acute and long-term outcome variables, including the occurrence of SH and DKA, the level of glycemic control as measured by glycosylated hemoglobin (HbA1c), and the risks for the development of microvascular complications. However, there have been significant advances in diabetes management since that time, such as the introduction of new insulin analogs; smaller, faster, and more accurate self-blood glucose monitoring (SBGM) systems; sophisticated continuous subcutaneous insulin infusion (CSII) pumps; and CGM, which may have changed outcomes associated with both conventional and intensive therapy. It is reasonable to assume that any new diabetes treatment should be capable of achieving outcomes that are significantly better than those achieved in either the conventional or intensive control groups in that study that began in 1993, nearly 30 years ago. However, reports from large type 1 diabetes mellitus (T1DM) databases that include children and adults using insulin twice daily, multiple daily injections (MDI), or CSII show that mean HbA1c levels remain significantly above the ADA recommended target of <7.0% (\(\text{Table 1}\)).\(^11-14\) One of these databases, the T1D Exchange Registry was begun in 2010 to collect prospectively demographic, treatment, and outcome information from a large number of persons with T1DM from across the United States. The only criterion for inclusion in this registry is that the individual must have had T1DM since diagnosis. With over 8000 registrants to date, it is a reasonable, contemporary data set of conventional and intensively controlled subjects that provides information regarding treatment outcomes.

**Determining Clinical Requirements**

Based on the foregoing discussion, it is proposed that the clinical requirements for CL control systems in terms of both safety and efficacy be at least equivalent (noninferior) to the safety and efficacy of available conventional therapy.
Hypoglycemia

The frequency of SH among persons with T1DM remains unacceptably high (Table 1).

Noninferiority safety criteria for CL systems should be based on well-documented SH frequencies in the general T1DM populations. It is possible that these reports may significantly underestimate the actual occurrence of SH events because patients may be reluctant to disclose these events that could affect driving privileges or employment opportunities. Exploratory use of CGM has revealed that large numbers of children and adults with T1DM using either MDI or CSII experience nocturnal hypoglycemia (<70 mg/dl) frequently and for prolonged periods of time and that they often ignore or sleep through CGM alarms that signal these events. The use of CGM systems alone has been associated with reductions in the frequency of SH, even when subjects are not given specific instructions on how to alter their management regimens.

In the Juvenile Diabetes Research Foundation (JDRF) randomized controlled trial of CGM vs SBGM in children and adults being treated with either CSII or MDI, SH events were too infrequent to demonstrate an effect of CGM (20 events per 100 patient years vs 26.3 events per 100 patient years). Similarly, in the Sensor-Augmented Pump Therapy for A1C Reduction (STAR 3) study, a comparison of sensor-augmented insulin pump vs sensor-augmented MDI therapy, no differences in the rates of SH were observed (13.31 events/100 patient years vs 13.48 events per 100 patient years). Subjects in both of these studies had baseline frequencies of SH that were much lower than those reported for intensively controlled subjects in the DCCT (62 events/100 patient years) and most of the large databases listed in Table 1. These carefully performed studies present contemporary SH frequency data that could be appropriate for use as goals for superiority safety and efficacy requirements for CL control systems, but not for noninferiority criteria. Beck and colleagues have suggested that the infrequency of the occurrence of SH in these studies suggests that very large numbers of subjects would need to be studied using a CL prototype to be able to demonstrate superiority. It is suggested that CGM-measured indices such as detection of hypoglycemic thresholds be used as an alternative to document safety and efficacy of these systems.

Diabetic Ketoacidosis

Although DKA is a frequent presentation of T1DM at onset, especially in children, its frequency in the years post diagnosis is relatively low (Table 1). In the TID Exchange Registry, in the 12-month period from September 2010 to August 2011, 5.7% of 4120 children (ages 11.9 + 3.6 years experienced DKA; 7.8 events per 100 patient years). In the JDRF CGM trial, there was one episode of DKA and in the STAR 3 study, five episodes. Three of the episodes (3/247 subjects) occurred in the sensor-augmented pump therapy group (2 adults, 1 child). Thus, a noninferiority safety criterion for CL control systems might be similar to the TID Exchange Registry data (7.8 events per 100 patient years), and a superiority safety criterion might be similar to that in the STAR 3 trial of sensor-augmented pumps (1.2 events per 100 patient years). However, the same argument regarding using the frequency of SH as a safety and efficacy outcome can be made for using episodes of DKA to document noninferiority or superiority, and it may be reasonable to use CGM-measured indices as an alternative to counting episodes of DKA. An additional

---

Table 1. Values for Hemoglobin A1c and Event Rates for Severe Hypoglycemia and DKA

<table>
<thead>
<tr>
<th>Study/author</th>
<th>Year</th>
<th>n</th>
<th>Value (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HbA1c</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hvidoere</td>
<td>2007</td>
<td>2100</td>
<td>8.2</td>
<td>11</td>
</tr>
<tr>
<td>SEARCH</td>
<td>2009</td>
<td>3947</td>
<td>8.18</td>
<td>12</td>
</tr>
<tr>
<td>Fritsch</td>
<td>2011</td>
<td>28,770</td>
<td>8.21</td>
<td>13</td>
</tr>
<tr>
<td>T1D Exchange</td>
<td>2011</td>
<td>3802</td>
<td>8.4</td>
<td>14</td>
</tr>
<tr>
<td><strong>Severe hypoglycemia (episodes/100 pt years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCCT</td>
<td>1993</td>
<td>711</td>
<td>62</td>
<td>10</td>
</tr>
<tr>
<td>Rewers</td>
<td>2002</td>
<td>1243</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>UK Hypoglycemia</td>
<td>2007</td>
<td>57</td>
<td>320</td>
<td>16</td>
</tr>
<tr>
<td>Donnelly</td>
<td>2005</td>
<td>94</td>
<td>115</td>
<td>17</td>
</tr>
<tr>
<td>EDIC conventional</td>
<td>2005</td>
<td>606</td>
<td>39.6</td>
<td>18</td>
</tr>
<tr>
<td>EDIC intensive</td>
<td>2005</td>
<td>620</td>
<td>48.4</td>
<td>18</td>
</tr>
<tr>
<td>Hvidoere</td>
<td>2007</td>
<td>2100</td>
<td>27</td>
<td>11</td>
</tr>
<tr>
<td>JDRF CGM</td>
<td>2011</td>
<td>436</td>
<td>17.9</td>
<td>19</td>
</tr>
<tr>
<td>O’Connell</td>
<td>2011</td>
<td>1683</td>
<td>5.6</td>
<td>20</td>
</tr>
<tr>
<td><strong>DKA (episodes/100 pt years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rewers</td>
<td>2002</td>
<td>1243</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>EDIC</td>
<td>2005</td>
<td>1226</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Hvidoere</td>
<td>2007</td>
<td>2100</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Karges</td>
<td>2010</td>
<td>10,682</td>
<td>5.2</td>
<td>19</td>
</tr>
<tr>
<td>Fritsch</td>
<td>2011</td>
<td>28,770</td>
<td>4.9</td>
<td>12</td>
</tr>
<tr>
<td>T1D Exchange</td>
<td>2011</td>
<td>4120</td>
<td>7.8</td>
<td>13</td>
</tr>
</tbody>
</table>
safety concern for CL control systems involves the unintended interruption of basal insulin delivery. Evidence from studies utilizing CSII suspension to prevent impending overnight hypoglycemia has demonstrated the safety of suspending basal insulin delivery for 90–165 min.3–5

**Glycemic Control Requirements**

It seems obvious that there should be some level of glycemic control as a clinical requirement for CL control systems. Deciding what measure of glycemic control to select may be problematic. Glycosylated hemoglobin is the obvious choice because most reports of glycemic control include HbA1c as an outcome variable. Recent reports have established the relationship between current laboratory and point-of-care assays and older DCCT standards and have demonstrated a decrease in the calibration variability among various laboratories.30 In addition, a large clinical study has established the relationship between HbA1c and average BG.31 The average HbA1c in the DCCT conventional treatment group was 9.4% (adolescents, 9.8%), while that in the intensive treatment group was 7.5% (adolescents, 8.1%).30,32 As presented in Table 1, average HbA1c in several large databases is approximately 8.2%. Thus, it seems appropriate that an HbA1c value no higher than 8.2% should be a noninferiority clinical criterion for CL control systems.

Superiority outcome criteria for HbA1c could be any HbA1c less than the noninferiority value of 8.2%. However, it is also reasonable to suggest that the HbA1c superiority criterion be the achievement of the American Diabetes Association’s HbA1c goals of 7.0% for adults, 7.5–8.5% in toddlers and preschoolers (ages 0–6 years), <8.0% in school-aged children (6–12 years), and <7.5% in adolescents (13–19 years).33 Such goals should be achievable in well-motivated compliant individuals given the results of the STAR 3 trial, in which the average HbA1c achieved was 7.3% in those over 19 years old and 7.9% in those 7–19 years old.27

**Figure 1** demonstrates the clinical problem associated with arbitrarily selecting an HbA1c level as a criterion for CL control systems. This figure represents the CGM tracings of two individuals with identical HbA1c values. It is apparent to even the casual observer that the glycemic control of these individuals differs significantly. While controversy still exists over the clinical significance of glycemic variability to short- and long-term health among T1DM individuals, it is clear that CGM has made it possible to observe this phenomenon to an extent unavailable previously.34–37 We will never be able to know the influence of variability on the results of the DCCT based on the data collected during that study.10 Given that glycemic variability is greater among persons with T1DM than among those with no diabetes, it seems reasonable that a goal of clinical care would be to reduce variability as much as would be feasible.36,37 Deciding how to analyze variability is not an easy task.38 It is important to note that standard deviation (SD) is not a recommended statistic to describe BG variability because the BG scale is asymmetric; BG values are not normally distributed. Therefore, SD is influenced more by hyperglycemia than by hypoglycemia. Statistics such as MAGE (mean amplitude of glycemic excursions), which are based on SD, are similarly insensitive to hypoglycemia. However, SD is an appropriate statistic for use when describing the rate of change of BG scale and the stability of CL control overtime.

Perhaps the simplest way to characterize variability is by recording time spent within a target range, for instance 70–180 mg/dl, as well as time spent below and above that range. Such computations may be tedious but sophisticated CGM data recorders can provide this information easily. The Diabetes Research in Children Network study group has demonstrated an increase in the percentage (52–60%) of BG values with the range of 70–180 mg/dl among children using a CGM system over 12 weeks.39 Others have shown a 21% reduction in time spent with BG <55 mg/dl, a 23% reduction in time spent with BG >240 mg/dl, and a 26% increase in time spent within a range of 81–140 mg/dl among adults using a CGM system for as short a time period as 3 days.24 In the JDRF CGM study, patients >25 years old reduced their mean min/day with BG >180 mg/dl while increasing their mean min/day with BG between 71 and 180 mg/dl.26 In the STAR 3 trial, adults and children reduced the area under the glucose curve measured when BG >180 mg/dl.27

More sophisticated statistics can be determined from CGM data that can provide important information regarding variability and risk assessment for both high and low BG.38 The blood glucose risk index (BGRI), the sum of the low blood glucose index (LBGI) plus the high blood glucose index (HBGI), provides a measure of the extent and frequency of BG fluctuations. Studies have demonstrated that the LBGI can predict 40–50% of the variance in the prediction of future low BG (BG < 70 mg/dl), while the HBGI correlates with postprandial BG levels and HbA1c. Visual interpretations of CGM data include histograms of rate of BG change and Poincaré plots of BG (f_{G,i}) vs
BG (\(t_i\)) which demonstrate the stability of the BG system. Control variability grid analysis permits graphical representation of minimum vs maximum BG levels over a designated period of time.\(^{40}\) It remains to be determined which of these representations of variability will emerge as a standard.

**Psychobehavioral Variables**

Two important psychobehavioral variables should be included when considering clinical requirements for a CL therapy system—Diabetes Quality of Life and Fear of Hypoglycemia.\(^{41-43}\) Each can be measured objectively with well-validated surveys. Diabetes Quality of Life was developed during the DCCT and consists of four subscales—life satisfaction, diabetes impact, worries about diabetes, and social/vocational concerns. The Hypoglycemia Fear Survey has two subscales—worry about hypoglycemia and behaviors to avoid hypoglycemia. Fear of hypoglycemia has been shown to be significantly related to a history of SH in adults and adolescents with T1DM, to be common among persons using insulin, and to be reduced during CGM use in the JDRF CGM study.\(^{42}\) It is conceivable that either of these psychobehavioral variables could be positively affected by a CL control system that does not improve an individual’s glycemic control. Such an outcome could be considered positive (or superior) for any one such individual.

---

**Figure 1.** Continuous glucose monitor profiles of two patients with T1DM and identical HbA1c values.
Variables Related to System Management

Patient management of insulin infusion by a CL system should be at least as easy and safe as the current reference therapies of T1DM (noninferiority) and at best reduce the burden associated with the clinical use of these treatments (superiority).

The first important requirement deals with the monitoring of the stability of insulin delivery related to the infusion catheter. Indeed, currently used insulin pumps are safe and reliable for long durations of several years. In contrast, infusion catheters request changes every 3–4 days because of the body’s reactions occurring at the subcutaneous infusion site. Moreover, 3–9% of inserted catheters perform defective delivery within 6 h of insertion and need to be replaced, and 8% of used catheters present an impaired infusion performance before their expected lifetime, as shown in a study.\(^\text{44}\)

In addition, the wear-time of catheters has a demonstrated influence on the pharmacokinetics of a short-acting insulin analog.\(^\text{45}\) In order to obtain the benefits of an algorithm-driven insulin therapy, the response of insulin infusion has to be kept stable. Because self-adjustment of insulin delivery rate is transferred from the patient to the system in CL use, it is of utmost important to warn the system and the patient that the intended insulin infusion has been altered. Safety systems include the detection of increased internal pressure in the catheter. Their reliability is often questioned in clinical practice, e.g., bending of a soft catheter cannula results in an intermittently increased pressure that can greatly impair insulin delivery. Besides an internal pressure alarm, a flow alarm will be an important additional requirement of CL insulin delivery. A compensation of the flow alteration by the system would be a first outcome because the patient will not self-act on the infusion rate. A request to change the catheter would be the ultimate outcome in case of failure of the system compensation. Although such alarms would also benefit all pump users, they are needed expressly in a CL insulin delivery system to prevent sustained erroneous delivery of insulin.

A second important requirement is the monitoring of glucose sensing accuracy. Despite continuous improvements, CGM data keep a mean average relative deviation between 10% and 20% vs paired capillary BG measurements.\(^\text{46}\) Iterative calibrations of the CGM signal against capillary BG level aim at minimizing the error in estimation of BG level. Because the output of CL systems in terms of insulin infusion is tightly related to inputs coming from CGM, the accuracy of glucose sensing is essential. Various options may be considered to obtain an efficient monitoring of sensor accuracy. In accordance with the indices of targeted glucose control mentioned earlier, a request for performing a capillary BG measurement each time sensor glucose exits the targeted range (e.g., 70–180 mg/dl) could be a possible option. Predefined HBGI and LBGI thresholds might be other more sophisticated alternatives, thanks to their online computations by the system. Retrospective timely fitting of sensor glucose levels from scheduled iterative calibration points has also been suggested.

Discussion

The U.S. FDA has recently released for public comment proposed guidelines for premarket approval applications for artificial pancreas device systems.\(^\text{47}\) These guidelines include recommendations for assessing the safety and effectiveness of these systems. The stated goal for these systems is to “maintain glucose values within range or near a target while minimizing adverse events such as hypo- and hyperglycemia.” In terms of safety, the CL system should not increase the incidence of SH or DKA. The FDA proposes using a CGM-based correlate for determining hypoglycemic events at a threshold of either 60 or 70 mg/dl. A similar CGM-based end point for hyperglycemia might be BG >240 mg/dl with ketonuria. In terms of efficacy, the FDA suggests that two superiority end points might be a reduction in mean HbA1c by 0.4% and a 30% reduction of either SH or CGM-based hypoglycemic events.

The guidelines suggest that noninferiority and superiority comparisons be made to standard-of-care therapy. Unfortunately, standard of care is defined as sensor-augmented pump therapy. We would strongly disagree with this definition as it clearly does not reflect the therapeutic regimens used to treat the majority of persons with T1DM. Such a definition could lead to labeling of CL systems that would restrict their use to persons who are already utilizing sophisticated state-of-the-art therapy to treat their diabetes and who may not realize a significant acute or long-term benefit from this therapy. Also, those who may benefit the most from CL therapy might lose the opportunity to reduce their risk of acute and long-term complications as well as improve their quality of life. Researchers and diabetes health care professionals should take the opportunity to familiarize themselves with these proposed guidelines and submit comments.
Clinical requirements for CL control systems suggested in this paper are summarized in Table 2. They are based on a careful review of contemporary safety and efficacy data associated with a variety of insulin regimens and include noninferiority and superiority criteria for important psychological variables such as quality of life and fear of hypoglycemia. Manageability criteria are proposed as well. While it is reasonable and efficient to utilize CGM-based documentation of hypo- and hyperglycemic events in relatively short length studies to satisfy safety and efficacy concerns for marketing CL systems for the population of persons with T1DM, consideration should be given to encouraging (requiring) industry-sponsored registries to track longitudinally variables of clinical concern, including frequencies of SH and DKA, and measures of glycemic control.

### Table 2. Clinical Criteria for Closed-Loop Control Systems

<table>
<thead>
<tr>
<th></th>
<th>Noninferiority</th>
<th>Superiority</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Safety</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypoglycemia</td>
<td>48.4/100 patient years</td>
<td>&lt;48.4/100 patient years</td>
</tr>
<tr>
<td>DKA</td>
<td>7.8/100 patient years</td>
<td>&lt;7.8/100 patient years</td>
</tr>
<tr>
<td><strong>Efficacy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycemic control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HbA1c</td>
<td>8.2%</td>
<td>&lt;8.2%</td>
</tr>
<tr>
<td>Glycemic variability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time in range</td>
<td>Current</td>
<td>Increase</td>
</tr>
<tr>
<td>BGR</td>
<td>Current</td>
<td>Decrease</td>
</tr>
<tr>
<td><strong>Manageability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reliability of infusion catheter</td>
<td>Current</td>
<td>Increase</td>
</tr>
<tr>
<td>Accuracy of glucose monitoring</td>
<td>Current</td>
<td>Increase</td>
</tr>
</tbody>
</table>

Disclosures:
Dr. Renard has received honoraria from Medtronic, Inc., Roche Diagnostics, Animas, Eli-Lilly, Novo-Nordisk, Sanofi-Aventis, and Abbott Diabetes Care for serving as consultant.

References:
Clinical Requirements for Closed-Loop Control Systems


