The promise of automated insulin delivery (AID), also called artificial pancreas (AP), has been demonstrated in clinical experiments since the 1970s (1). The main challenge was to advance the technology from proof-of-concept experiments in a clinical environment to daily use under free-living conditions. Progress in hardware and control technologies enabled the development of various systems during the past decade, culminating in the commercial availability of the first hybrid closed-loop AP in 2018 (2). Worldwide research and development in industry and universities contributed to the development of alternative technologies, and fully automated AP systems will be introduced in the near future.

The AP automates the information collection, decision-making, and insulin management of a person with type 1 diabetes to maintain euglycemia despite various daily disturbances such as meals, physical activity, sleep, and stress. All AP systems have three basic components: sensors, decision-making algorithms (controllers), and insulin infusion pumps. The first AP systems were conceived to automate the activities and decisions of the patient, and the AP consisted of a glucose concentration sensor and transmitter, a control algorithm, and an insulin pump. Over the years, the sizes of the devices have been reduced, and the current commercial APs and alternative designs that are in the final stages of clinical trials consist of a continuous glucose monitoring (CGM) and transmitter system that communicates with a control unit housed in an insulin pump. All devices and algorithms are reviewed by the U.S. Food and Drug Administration and cleared after a rigorous evaluation before becoming available to the public.

Reports on the first clinical experiments of closed-loop AID date back to 1974, with intravenous administration of dextrose to simulate food, intravenous sampling of blood glucose concentration (BGC), and intravenous infusion of insulin (1). The closed-loop control of BGC was a remarkable success, regulating BGC much better than the multiple daily injections performed by the subject. It took more than four decades to move AP technology from the hospital room to daily living at home (3–10).
An important challenge in this transition was the time delay in getting automated glucose concentration readings at high frequency and delivering insulin to the bloodstream under free-living conditions. Patient safety and security concerns dictate that there should be no direct open access to vasculature for glucose measurements or insulin infusion when the AP will be used in daily living. Glucose concentration in interstitial fluid is used to estimate BGC, and insulin is delivered to subcutaneous tissue to satisfy these safety constraints. Hence, the glucose concentration reported by CGM systems provides delayed information about BGC, and the diffusion of the infused insulin from the subcutaneous tissue to the bloodstream delays insulin action. When multiple daily injections are used for insulin delivery, the delay in insulin reaching the bloodstream is compensated for by adjusting the insulin injection time (e.g., giving an injection 15–30 minutes before a meal). In the first generation of APs, manual announcement of meals to the AP compensates for the mass transfer delay of insulin boluses delivered by the AP. Similarly, manual intervention reduces or shuts down insulin infusion before beginning a typical bout of medium-intensity exercise. Such manual inputs of information are concessions to full automation of the AP. The available hybrid closed-loop AP (Medtronic Minimed 670G) uses such manual information entry. Some of the other investigational APs have relaxed the manual meal announcement, but manual adjustment of insulin infusion dose before exercise is still used in these systems. Advances in modeling and control techniques and new wearable devices that provide real-time streaming data provide new opportunities to advance AP systems and accommodate many daily disturbances in a fully automated manner.

Advances in glucose sensor technologies have increased sensor accuracy, extended sensor use life, and eliminated calibrations with fingerstick blood glucose measurement (11). Insulin pump technologies and delivery options have also improved. The strategic initiatives of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and JDRF have enabled many of these technological advances and facilitated a fertile environment for AP development. JDRF initiated an AP research consortium and proposed a six-step sequential roadmap to AP system development (12).

The initial JDRF roadmap proposed the sequential development of 1) an insulin pump with a shut-off function when hypoglycemia was determined from CGM readings, 2) use of BGC predictions to trigger the insulin shut-off function, 3) a hypoglycemia/hyperglycemia minimizer that predicts BGC and reduces or shuts off insulin infusion if hypoglycemia is predicted or provides a bolus if hyperglycemia is predicted, 4) an automated basal/hybrid closed-loop AP with manual meal inputs, 5) a fully automated closed-loop AP, and finally 6) a fully automated multihormone closed-loop AP. The first four milestones have been accomplished, and various systems offering these features are available. The JDRF roadmap was recently updated based on achievements in AP research and advances in technology since 2009, condensing the AP development roadmap to three steps and bifurcating automated delivery approaches to either insulin-only or multihormone systems (using insulin and glucagon, insulin and amylin, or insulin and other glucose-modulating agents) (13).

Control algorithms with significantly different philosophies have been proposed for AP systems. Three different types of algorithms—proportional-integral-derivative (PID) control, model predictive control (MPC), and fuzzy-logic knowledge-based systems—represent these philosophies in automating insulin delivery in AP systems (Table 1).

**PID Control Algorithms**

PID control systems have been used in various industries since the 1940s. They compute the control action based on the difference (called the error) between the reference (or desired) BGC and the measured value of BGC. This error is processed in three different ways: the proportional term considers the current value of the error, the integral term considers the sum of the errors over a past time window, and the derivative term considers the rate of change in the last two errors. Each term is multiplied by a coefficient to adjust its contribution to the computation of the amount of insulin to be infused by the pump. These three coefficients are the adjustable parameters of the controller to make it more aggressive or more conservative. The proportional and derivative actions are also similar to the way the pancreas reacts to an increase in BGC similar to the first-phase insulin response, by releasing the insulin already available in the pancreas, and then producing and continuing to release insulin while the rate of change in BGC increase persists, similar to the second-phase insulin response. The integral action ensures that insulin infusion will continue as long as the error persists.

The simplicity of the PID logic and the similarity of the proportional and derivative actions to the behavior of pancreatic β-cells have been the main appeal of the PID approach. But this simple structure is limited in its ability to provide effective control when faced with the large spectrum of disturbances affecting the BGC and the continuous changes in metabolism as a function of many factors. The PID algorithm has been modified, and many auxiliary modules have been added to it to enhance its performance (14–16). The control algorithm in the Medtronic Minimed 670G is based on the PID approach.

**MPC Algorithms**

Most AP systems under development or in clinical trials use some variant of
MPC algorithms. Model-based control techniques use a dynamic model of glucose and insulin dynamics to predict how the BGC will vary in the future in response to a hypothetical set of future insulin infusions and to minimize the difference between the future BGC reference trajectories and BGC estimated by the model. MPC algorithms have four key elements: 1) a dynamic model of glucose and insulin dynamics for predicting future BGC values, 2) an “objective function” that includes the sum of the future errors between future BGC reference trajectories and BGC estimated by the model and the sum of the future insulin consumptions, 3) an optimization algorithm to minimize the objective function defined, and 4) constraints on the values and rates of change of BGC and insulin. Hence, the optimal future insulin infusions are computed by the algorithm based on BGC predictions for various hypothetical insulin infusion scenarios.

BGC prediction may be based on compartmental models describing physiological changes or data-driven models that use the current and recent past values of CGM readings and insulin infusion doses. The parameters of these models are computed by using historical data sets. Some early AP algorithms used an average model for all patients, but today the model parameters are personalized. These parameters may be fixed for each patient or adapted over time to the current BGC dynamics of the user. One set of AP algorithms adapt the parameters once daily based on recent data (mostly data from the previous day) (21), whereas others use recursive identification, through which the parameters may be updated with every new CGM data point reported (5-minute sampling time) (18,19).

Most AP algorithms use CGM and insulin pump data and, if entered, manually announced meal information in their BGC prediction models. New multivariable AP systems also use information captured from wearable devices (i.e., wristbands) such as heart rate, energy expenditure, and galvanic skin response to enhance the BGC prediction accuracy during periods of physical activity (22,23).

The objective function expresses the desired outcomes of the AP activities in regulating the BGC. It has two terms: the sum of the future errors between future BGC reference trajectories and BGC estimated by the model, and the sum of the future insulin infusion doses. The parameters of these models are computed by using historical data sets. Some early AP algorithms used an average model for all patients, but today the model parameters are personalized. These parameters may be fixed for each patient or adapted over time to the current BGC dynamics of the user. One set of AP algorithms adapt the parameters once daily based on recent data (mostly data from the previous day) (21), whereas others use recursive identification, through which the parameters may be updated with every new CGM data point reported (5-minute sampling time) (18,19).

### TABLE 1. Comparison of MPC, PID, and Fuzzy-Logic Algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>PID</td>
<td>• Used successfully in various industries since the 1940s</td>
<td>• The native algorithm uses only the current difference between measured (CGM) and desired BGC</td>
</tr>
<tr>
<td></td>
<td>• Good for single-input (CGM), single-output (insulin dose) systems with simpler dynamics</td>
<td>• Many additional modules added to predict BGC trends and to accommodate constraints</td>
</tr>
<tr>
<td></td>
<td>• Easy-to-understand algorithm with three terms (proportional, integral, and derivative)</td>
<td>• Constraints usually restrict the insulin dose computed by the PID algorithm</td>
</tr>
<tr>
<td></td>
<td>• Low computational burden</td>
<td>• Difficult to personalize the controller parameters</td>
</tr>
<tr>
<td>MPC</td>
<td>• Used successfully in various industries since the 1970s</td>
<td>• Based on advanced concepts in systems and control engineering</td>
</tr>
<tr>
<td></td>
<td>• The model can predict future BGC and optimize future insulin infusion doses in real time</td>
<td>• Higher computational burden caused by real-time optimization</td>
</tr>
<tr>
<td></td>
<td>• Accommodates constraints on BGC and insulin infusion during optimization</td>
<td>• Relies on availability of mathematical models describing glucose-insulin dynamics</td>
</tr>
<tr>
<td></td>
<td>• Can use adaptive, multivariable, and/or nonlinear models for better BGC prediction accuracy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Can use adaptive controllers to personalize the MPC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Can use multivariable models and controllers to mitigate the effects of physical activity and stress</td>
<td></td>
</tr>
<tr>
<td>Fuzzy-logic knowledge-based systems</td>
<td>• Expertise of the physician in treatment is captured and used in rule development; easy to understand by HCPs</td>
<td>• High cost of maintenance and revision of the rule base</td>
</tr>
<tr>
<td></td>
<td>• Fuzzy-logic accommodates day-to-day variations and disturbances</td>
<td>• As the number of rules increases, conflict resolution algorithms are needed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Medical and programming expertise is needed for personalizing the system</td>
</tr>
</tbody>
</table>

**MPC** algorithms. Model-based control techniques use a dynamic model of glucose and insulin dynamics to predict how the BGC will vary in the future in response to a hypothetical set of future insulin infusions and to minimize the difference between the future BGC reference trajectories and BGC estimated by the model (5,17–20). MPC algorithms have four key elements: 1) a dynamic model of glucose and insulin dynamics for predicting how the BGC will vary in the future in response to a hypothetical set of future insulin infusions and to minimize the difference between the future BGC reference trajectories and BGC estimated by the model (5,17–20). MPC algorithms have four key elements:

1. A dynamic model of glucose and insulin dynamics for predicting how the BGC will vary in the future in response to a hypothetical set of future insulin infusions.
2. An “objective function” that includes the sum of the future errors between future BGC reference trajectories and BGC estimated by the model and the sum of the future insulin consumptions.
3. An optimization algorithm to minimize the objective function defined.
4. Constraints on the values and rates of change of BGC and insulin. Hence, the optimal future insulin infusions are computed by the algorithm based on BGC predictions for various hypothetical insulin infusion scenarios.

BGC prediction may be based on compartmental models describing physiological changes or data-driven models that use the current and recent past values of CGM readings and insulin infusion doses. The parameters of these models are computed by using historical data sets. Some early AP algorithms used an average model for all patients, but today the model parameters are personalized. These parameters may be fixed for each patient or adapted over time to the current BGC dynamics of the user. One set of AP algorithms adapt the parameters once daily based on recent data (mostly data from the previous day) (21), whereas others use recursive identification, through which the parameters may be updated with every new CGM data point reported (5-minute sampling time) (18,19).

Most AP algorithms use CGM and insulin pump data and, if entered, manually announced meal information in their BGC prediction models. New multivariable AP systems also use information captured from wearable devices (i.e., wristbands) such as heart rate, energy expenditure, and galvanic skin response to enhance the BGC prediction accuracy during periods of physical activity (22,23).

The objective function expresses the desired outcomes of the AP activities in regulating the BGC. It has two terms: the sum of the future errors between future BGC reference trajectories and BGC estimated by the model, and the sum of the future insulin infusion doses. The parameters of these models are computed by using historical data sets. Some early AP algorithms used an average model for all patients, but today the model parameters are personalized. These parameters may be fixed for each patient or adapted over time to the current BGC dynamics of the user. One set of AP algorithms adapt the parameters once daily based on recent data (mostly data from the previous day) (21), whereas others use recursive identification, through which the parameters may be updated with every new CGM data point reported (5-minute sampling time) (18,19).
given sequence of hypothetical future insulin infusion doses; Figure 1), 2) control command (insulin infusion dose) sequence horizon (the future time window for hypothetical insulin infusion doses that are optimized), 3) penalty weights on the differences between BGC reference trajectory \(r\) and the BGC estimates \(\hat{y}\) where the circumflex [“hat” symbol] indicates that \(y\) is estimated) for all values in the prediction horizon, and 4) the penalty weights on the future insulin infusion doses. The values assigned to these weights declare the relative importance of the magnitudes of the differences between \(r\) and \(\hat{y}\) for every future BGC estimate in the prediction horizon and the cost of insulin use in every dose, respectively. For example, a larger deviation of BGC from the desired value may be tolerated for a short period of time during a meal or exercise by assigning lower weights at those time instants, and the system may be forced to reduce the deviation as time progresses by assigning larger weights for subsequent times. Furthermore, the relative magnitudes of the penalty weights between BGC deviations and insulin usage are used to adjust the relative importance of these two factors.

Figure 1 illustrates the real-time optimization of the objective function to regulate BGC. A set of insulin infusion values is suggested (denoted as the first sequence, blue dashed line) at the current time \(k\) so that BGC values inside the prediction horizon \((p)\) are computed by using the BGC prediction model. These predicted BGC values are denoted as first sequence of estimated BGC (blue filled circles connected by blue line). The BGC estimates diverge from the reference BGC trajectory (red line). Hence, the algorithm suggests a second set of insulin infusion values (magenta dotted line). BGC values inside the prediction horizon \((p)\) are computed again by using the BGC prediction model denoted as second sequence of estimated BGC (magenta filled circles connected by magenta line).

This infusion dose sequence yields much closer values of BGC estimates to the reference trajectory. These adjustments in insulin infusion dose sequence followed by BGC estimation and the evaluation of the objective function continue until the successive values of the objective function indicate that the minimum value is reached. The optimization algorithm automates the search for the optimal sequence of insulin infusion values to minimize the objective function. The first insulin infusion value (at time \(k\)) of this optimal sequence is sent to the insulin pump, and the whole optimization activity is repeated when the next CGM reading is transmitted to the AP (at time \(k+1\)). Since the number of insulin infusion values that will be computed affects the duration of the optimization, the control horizon (denoted by \(m\) in Figure 1) is smaller than the prediction horizon (denoted by \(p\)) to reduce the computational burden. In AP algorithms, \(p\) is usually 24 for a 2-hour prediction or higher, and \(m\) can be in the range of 4–8 (rather than 24) to reduce the computation time in the optimization. The remaining values of the insulin infusion dose (from \(k+m-1\) to \(k+p\)) are set equal to the insulin infusion dose at time \(k+m-1\).

Another advantage of MPC algorithms is the ability to impose constraints on the range of BGC, maximum insulin infusion dose, maximum change in two successive insulin doses, or other factors that the system developer has identified for limitation. Then, the optimization takes these constraints into account and does not provide a solution that would violate these constraints.

**Fuzzy-Logic Knowledge-Based Systems**

Knowledge-based systems capture the expertise of a care provider and the specific characteristics of an individual with type 1 diabetes in the form of “if-then” rules. Inferences are made by executing these rules, and insulin infusion suggestions are made based on the current state of the person (i.e., CGM data, meal information, recent doses of insulin infusions, historical data, and demographic information). Fuzzy-logic is used to accommodate the day-to-day variations in unmeasured disturbances such as spontaneous physical activity and the occurrence of stressful events. APs based on fuzzy-logic knowledge-based systems have been tested in clinical trials, and successful outcomes compared to CGM plus multiple daily injections or continuous subcutaneous insulin infusion have been reported (6).

One disadvantage of this approach is the high cost of maintenance of the system and the level of effort needed for modification to each patient. As
the number of rules in the rule base increases, some rules may conflict with others, and conflict resolution schemes are also needed.

**Multivariable AP Control Algorithms**

Advances in wearable devices have ushered in new paradigms in AID to people with type 1 diabetes. The first generation of AP systems relies only on CGM and pump data and manual inputs of additional information. Hence, the AP reacts to the change in BGC and infrequent manual inputs on major events that affect BGC (i.e., meals and the start of exercise). The integration of data from CGM systems and physiological data from wristbands in real-time provides new powerful alternatives for the management of type 1 diabetes.

The CGM-only AP is basically a reactive device (except for manual inputs of meals or start of exercise), since no additional metabolic or physiological information is used to anticipate future changes in BGC and enable proactive measures to regulate it. This operation is different from a normally functioning pancreas that rapidly detects changes in BGC and other hormones to make insulin administration decisions. APs do not have information about the other hormones (no sensors are available for use in free-living environments and with real-time reporting) and have delays in detecting BGC changes and larger delays in diffusion of infused insulin. The multivariable AP (mAP) approach includes proactive decisions based on real-time physiological information from wearable devices that reduce the delays in decision-making to adjust insulin sooner (8,19,25). The mAP approach merges the information from CGM (feedback information) and wearable devices (feed-forward predictive information) and provides a viable solution to develop the next generation mAPs that can better maintain BGC in a desired range in response to various challenges of daily living.

The mAP approach has several modules (Figure 2). Hypoglycemia prediction (30 minutes ahead of a potential hypoglycemic episode) is similar in concept to predictive hypoglycemia alarm systems found in most APs. But it uses a multivariable model to estimate BGC that includes inputs from a wristband (energy expenditure and galvanic skin response) to improve the prediction accuracy, especially during physical activities (19). The hyperglycemia prediction module extracts meal information (carbohydrate content estimates) from CGM data and provides timely information for computation of insulin boluses (26,27). The exercise assessment module detects the presence, type, intensity, and duration of physical activity and provides refined information to BGC estimation and control algorithms (23). The plasma insulin estimation module uses CGM and insulin pump data along with the demographic information of the user to estimate the plasma insulin concentration in real time (28). These estimates are used in the BGC estimation and control algorithms instead of the generic insulin-on-board information. Data from wearable devices also provide valuable information on the occurrence and characteristics of acute psychological stress and on sleep characteristics. This information is used to predict their effects on insulin sensitivity and improve predictions of BGC variations and the accuracy of control decisions.

AP algorithms can also be used by people with diabetes who do not use insulin pumps. The algorithms can provide suggestions/recommendations to people who prefer to use a pump with manual adjustments, a smart pen/cap, or multiple daily insulin injections. In this case, the AP algorithms will be embedded in a decision-support system to provide insulin dose suggestions and recommendations on meals and physical activity and leave the decision for the final action to the user.

**Dual-Hormone AP Systems**

Dual-hormone AP systems mimic the biological pancreas more realistically by administering both insulin and glucagon in response to changes in BGC. In dual-hormone APs, glucagon is given as mini-boluses to prevent or treat hypoglycemia while insulin infusion is suspended (20,29–31). The same types of control algorithms can be used for glucagon control. The presence of two control loops may cause chattering between the two loops, but this is prevented in the multihormone AP, since only one of...
the loops is active to deliver only one of the hormones at any given time.

Conclusion
Various alternatives in AP control algorithms provide a strong foundation for the development of fully automated AP control systems. Four separate projects funded by the NIDDK are testing fully automated APs using CGM systems, and the successful outcomes will enable requests of regulatory approval for use (32). Dual-hormone and multivariable AP systems offer additional capabilities that will improve BGC regulation more tightly and further enhance the quality of life of people with type 1 diabetes.

Acknowledgments
Multivariable AP research is supported by the National Institutes of Health under grants 1DP3DK101077-01 and 1DP3DK101075-01 and by JDRF under grants 17-2013-472 and 2-SRA-2017-506-M-B.

Duality of Interest
No potential conflicts of interest relevant to this article were reported.

Author Contributions
A.C. is the sole contributor to and guarantor of this work and, as such, had full access to all the data and concepts developed in the study and takes responsibility for the integrity of the information.

References