Outcome Measures for Artificial Pancreas Clinical Trials: A Consensus Report

David M. Maahs,[1] Bruce A. Buckingham,[2] Jessica R. Castle,[3] Ali Cinar,[4] Edward R. Damiano,[5] Eyal Dassau,[6,7] J. Hans DeVries,[8] Francis J. Doyle III,[6,7] Steven C. Griffen,[9] Ahmad Haidar,[10,11] Lutz Heinemann,[12] Roman Hovorka,[13] Timothy W. Jones,[14] Craig Kollman,[15] Boris Kovatchev,[16] Brian L. Levy,[17] Revital Nimri,[18] David N. O'Neal,[19] Moshe Philip,[18] Eric Renard,[20] Steven J. Russell,[5,21] Stuart A. Weinzimer,[22] Howard Zisser,[6] and John W. Lum[15]

Research on and commercial development of the artificial pancreas (AP) continue to progress rapidly, and the AP promises to become a part of clinical care. In this report, members of the JDRF Artificial Pancreas Project Consortium in collaboration with the wider AP community 1) advocate for the use of continuous glucose monitoring glucose metrics as outcome measures in AP trials, in addition to HbA$_{1c}$, and 2) identify a short set of basic, easily interpreted outcome measures to be reported in AP studies whenever feasible. Consensus on a broader range of measures remains challenging; therefore, reporting of additional metrics is encouraged as appropriate for individual AP studies or study groups. Greater consistency in reporting of basic outcome measures may facilitate the interpretation of study results by investigators, regulatory bodies, health care providers, payers, and patients themselves, thereby accelerating the widespread adoption of AP technology to improve the lives of people with type 1 diabetes.

Since the publication of the Diabetes Control and Complications Trial (DCCT) in 1993 (1), the main outcome measures for glycemic control in people with type 1 diabetes have been hemoglobin A1c (HbA$_{1c}$) due to the clear link to the development of complications and episodes of severe hypoglycemia (SH) as it is an immediate life-threatening event. Advances in diabetes treatment and technology have since resulted in improved care, reflected in lower HbA$_{1c}$ and rates of SH for people with type 1 diabetes (2–6). However, many patients still struggle with glucose control and have large and erratic swings in glycemia (7). Research on and commercial development of the artificial pancreas (AP), either as automated insulin-only delivery or as multihormonal delivery, continue to progress rapidly, and the AP promises to become a part of clinical care (8). An AP system may benefit individual patients in unique ways that would not be reflected in HbA$_{1c}$ improvements alone; for example, a patient with a low HbA$_{1c}$ and frequent hypoglycemia may have an increase in HbA$_{1c}$ on an AP system while hypoglycemia and quality of life improve.

OBJECTIVE AND RATIONALE

In this report, members of the JDRF Artificial Pancreas Project Consortium in collaboration with the wider AP community 1) advocate for the use of continuous glucose monitoring (CGM) glucose metrics as glycemic outcome measures in AP trials, in addition to HbA$_{1c}$, and 2) identify a short set of basic, easily interpreted outcome measures to be reported in AP studies whenever feasible. Currently, the U.S. Food and Drug Administration accepts the use of various CGM glucose metrics in AP trials (9), but investigators do not always use a consistent set of measures that enables comparison. Thus, one rationale for the current report is to enable basic comparison between different AP research studies and with other clinical studies on

[1]Barbara Davis Center for Childhood Diabetes, University of Colorado School of Medicine, Aurora, CO
[2]Division of Pediatric Endocrinology and Diabetes, Department of Pediatric Endocrinology, Stanford University, Stanford, CA
[3]Department of Medicine, Division of Endocrinology, Harold Schnitzer Diabetes Health Center, Oregon Health & Science University, Portland, OR
[4]Department of Chemical and Biological Engineering, Illinois Institute of Technology, Chicago, IL
[5]Department of Biomedical Engineering, Boston University, Boston, MA
[6]Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, CA
[7]Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA
[8]Department of Endocrinology, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands
[9]JDRF, New York, NY
[10]Biomedical Engineering Department, McGill University, Montreal, Quebec, Canada
[11]Division of Endocrinology, McGill University, Montreal, Quebec, Canada
[12]Profil Institute for Metabolic Research, Neuss, Germany
[13]Wellcome Trust-Medical Research Clinical Institute of Metabolic Science, University of Cambridge, Cambridge, U.K.
[14]Telethon Kids Institute, The University of Western Australia, Perth, Australia
[15]Jaeb Center for Health Research, Tampa, FL
[16]University of Virginia, Charlottesville, VA
[17]Johnson & Johnson Diabetes Care Companies, Wayne, PA
[18]Jesse Z and Sara Lea Shafer Institute of Endocrinology and Diabetes, National Center for Childhood Diabetes, Schneider Children's Medical Center of Israel, Sackler Faculty of Medicine, Tel Aviv University, Petah Tikva, Israel
[19]Department of Medicine, University of Melbourne, St. Vincent's Hospital Melbourne, Fitzroy, Victoria, Australia

glycemic control in type 1 diabetes. We acknowledge there are methodological limitations with between-study comparison that require careful consideration of study design differences. However, the standardization of these simple metrics provides a starting point for regulators, payers, health care providers, and patients to interpret AP and other study data with interventions on glycemic control. This will be especially important as AP systems become part of the daily lives of people with type 1 diabetes. Standardization of these measures does not preclude the addition of other metrics specific for a particular AP approach or used by particular research groups. In this report, we specifically advocate for the use of a basic set of CGM glucose metrics in AP studies. We suggest that their use in general type 1 diabetes studies is broadly applicable and highly relevant given the increasing adoption of CGM in research and clinical care (7).

Improvements in and adoption of AP-related technology, particularly in the reliability and accuracy of CGM systems, have focused attention on determining the best metrics for assessing outcomes in studies with people with type 1 diabetes (10–16). Currently available CGM systems with glucose readings up to every 5 min, or 288 times daily, provide considerably more data than do the American Diabetes Association recommendation of checking blood glucose 6–10 times daily (17) or the 7-point blood glucose measurements performed quarterly for research purposes in the DCCT. Although 7-point blood glucose profiles do provide insights into glycemic excursions that are not apparent with HbA_{1c}, the profiles are very dependent on patient motivation and the chosen day of performance and provide only limited information about glucose control compared with glucose values provided by CGM systems. From a patient perspective, a CGM glycemic profile is more meaningful in that it shows highs, lows, trends, and variability as well as the effect of behaviors on

Table 1—Recommended basic outcome measures to be reported for AP clinical trials

	Comments
Glycemic metrics[*],[†]	
HbA_{1c}	If intervention period $\geq$3 months
Mean CGM glucose	
% CGM time <50 mg/dL (<2.8 mmol/L)	
% CGM time <60 mg/dL (<3.3 mmol/L)	
% CGM time <70 mg/dL (<3.9 mmol/L)	
% CGM time 70–140 mg/dL (3.9–7.8 mmol/L)	
% CGM time 70–180 mg/dL (3.9–10.0 mmol/L)	
% CGM time >180 mg/dL (>10.0 mmol/L)	
% CGM time >250 mg/dL (>13.9 mmol/L)	
% CGM time >300 mg/dL (>16.7 mmol/L)	
SD and coefficient of variation of CGM values	SD is much more dependent on the mean than coefficient of variation
Fasting blood glucose, mg/dL (mmol/L)	If available, depending on study design; CGM glucose at 06:00 can be taken as proxy
Safety metrics	
SH events	As defined by ADA (adults) (32) and ISPAD (children and adolescents) (31)
Diabetic ketoacidosis events	Per ADA definition (41)
Technical performance metrics[*]	
% Time closed-loop active	
Total daily dose of insulin	
Total daily dose of glucagon or other hormones	If applicable

ADA, American Diabetes Association; ISPAD, International Society for Pediatric and Adolescent Diabetes. [*]Metrics may have a skewed distribution. Report median (quartiles) instead of mean if not normally distributed. [†]All CGM measures should be reported for the overall 24-h period (if applicable) and also stratified by daytime and nighttime periods. The time period 00:00 to 06:00 is proposed as a definition of the nighttime period to exclude postprandial data as much as possible for a typical study population, though this definition may not be appropriate for all studies.

glucose levels (18–21). This contrasts with HbA_{1c} as a metric of integrated glycemic exposure over time. In particular, HbA_{1c} does not provide information on frequency and extent of hypo- or hyperglycemia, which is a crucial aspect to evaluating glucose control in people with type 1 diabetes.

BASIC OUTCOME MEASURES

The recommended basic set of outcome measures presented in Table 1 includes CGM glucose metrics to define time spent in desired ranges as well as time in hypo- and hyperglycemia, measures of CGM glucose variability, safety measures such as SH and diabetic ketoacidosis, and technical metrics to evaluate AP system performance. It is intended

that these measures be applicable across a wide range of AP study designs, including both short-term pilot studies and longer-term in-home or pivotal studies. Many of the glucose cut points and ranges are based on convention in AP research but were chosen to allow for comparison between studies.

HbA_{1c} remains the best currently available measure to assess long-term glycemic control and should be assessed in any AP study of 3 months or longer. However, it is clear that HbA_{1c} only captures average glycemia and does not provide information on the frequency or severity of hypoglycemic events. Although HbA_{1c} is currently the most accepted metric for risk stratification of long-term complications of diabetes,

[20]Department of Endocrinology, Diabetes, and Nutrition, Montpellier University Hospital, INSERM Clinical Investigation Centre 1411, Institute of Functional Genomics, CNRS UMR 5203, INSERM U1191, University of Montpellier, Montpellier, France
[21]Diabetes Unit and Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA

[22]Yale University School of Medicine, New Haven, CT

Corresponding author: John W. Lum, jlum@ jaeb.org.

the proposed metrics more comprehensively describe glycemia.

ADDITIONAL RECOMMENDATIONS AND LIMITATIONS

Graphical presentation of outcome data is also important, although standardization is less straightforward than with tabular data. A common figure for visualization of pooled-subject AP performance is the modal day glycemic control plot (or analogous insulin delivery plot) with median line and interquartile range bands. Inclusion of a cumulative histogram of CGM data would support the extraction of arbitrary glycemic ranges for comparison purposes (22). Numerous other graphical representations of data have been developed, and the choice of figures should be individualized for the data and the target audience (23).

The number of symptomatic hypoglycemia events per week may also be valuable as a meaningful clinical index of diabetes burden to the patient (24–27). Indeed, time spent below targeted glucose range according to CGM data may not fully capture the patient's experience with debilitating glucose-related events, which might better illustrate diabetes burden. Because reliable capture of symptomatic hypoglycemia events requiring treatment may be challenging in longer-term AP studies, biochemical hypoglycemia event rate as measured by CGM could be reported as a proxy. For example, the rate of CGM excursions below 70 or 55 mg/dL (3.9 or 3.0 mg/dL) for at least 10 or 30 min or longer time periods could be reported, as could other metrics including area under the curve (28–32). Many other novel measures of AP performance and algorithms have been developed (33–36), and this is an area of active research.

CGM and pump make and model and the kind of device running the control algorithm (e.g., laptop, smartphone) should be specified, including any relevant CGM signal conditioning algorithm details. We note that bias can occur when the same CGM that informs the AP controller is also used to assess glycemic outcomes (37), but there often is no practical alternative to this approach. Any special system- or protocol-related design elements should be disclosed if they are intended to improve safety or impact glycemic control or if they place additional burden on the user. The amount and timing of contact between study staff and participants in both the AP and comparator arms of the studies should be reported for in-home studies. In addition, CGM calibration logistics should be disclosed, along with a description of how conventional capillary blood glucose measurements are performed. Determination of median (or mean) absolute relative difference (MARD) for CGM versus capillary blood glucose is often used to characterize CGM accuracy in AP studies, though blood glucose sampling bias may limit the generalizability of these results (38).

Depending on the study design, the outcomes described could be reported for the entire cohort of a study or the study participants could be stratified into relevant subgroups with outcomes reported separately. For example, improved HbA_{1c} without increased risk for hypoglycemia could be reported for those who were poorly controlled at the baseline (e.g., baseline HbA_{1c} >8% [>64 mmol/mol]), whereas reduced incidence of hypoglycemia without deterioration in HbA_{1c} could be reported for those with well-controlled average glycemia at the baseline. The analysis of the primary and other important outcomes should be performed on an intention-to-treat basis.

Future areas of need for AP technology include expanded standardized metrics to evaluate the technical performance of AP systems (22) and to assess patient/caregiver usability, including psychosocial metrics such as quality of life and other measures of reduction of burden, which need to be developed, including stress, anxiety, depression, and use during exercise (39). In addition, the development of standard measures to assess patient preference may be used to support regulatory approval and to serve to inform health care providers and patients of the potential impact of the use of AP systems (40). Compelling health economic measures comparing AP system costs with the potential short- and long-term economic benefits are required to establish the financial viability of these systems and to drive acceptance by health care providers and people with diabetes (8).

Multiple large and longer in-home clinical trials will soon begin with different AP systems supported by the National Institute of Diabetes and Digestive and Kidney Diseases, JDRF, the Helmsley Charitable Trust, and other funders, as well as those being supported by the industry. Some of these have been designed as pivotal trials to provide data about relevant end points to be presented to U.S. Food and Drug Administration and other regulatory authorities for approval of AP products for clinical use and to support reimbursement. This report emphasizes the need for a set of basic, uniform, standardized, and comparable outcome measures with different AP systems. As AP technologies become available for clinical use, common data reports that compare systems will be desired for health care providers, payers, and people with type 1 diabetes and their families.

In summary, members of the JDRF Artificial Pancreas Project Consortium and the larger AP community advocate for the adoption of a set of basic outcome metrics that will allow for comparison between studies with different AP systems. This, in turn, will facilitate the interpretation of the information from trials and contribute to the ultimate goal of widespread adoption of AP technology to improve the life of people with type 1 diabetes.

References

1. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993;329:977–986

2. Bergenstal RM, Tamborlane WV, Ahmann A, et al.; STAR 3 Study Group. Effectiveness of sensor-augmented insulin-pump therapy in type 1 diabetes. N Engl J Med 2010;363:311–320

3. Hermanides J, Nørgaard K, Bruttomesso D, et al. Sensor-augmented pump therapy lowers HbA(1c) in suboptimally controlled Type 1 diabetes; a randomized controlled trial. Diabet Med 2011;28:1158–1167

4. Misso ML, Egberts KJ, Page M, O'Connor D, Shaw J. Continuous subcutaneous insulin infusion (CSII) versus multiple insulin injections for type 1 diabetes mellitus. Cochrane Database Syst Rev 2010;1:CD005103

5. Monami M, Lamanna C, Marchionni N, Mannucci E. Continuous subcutaneous insulin infusion versus multiple daily insulin injections in type 1 diabetes: a meta-analysis. Acta Diabetol 2010;47(Suppl. 1):77–81

6. O'Connell SM, Cooper MN, Bulsara MK, Davis EA, Jones TW. Reducing rates of severe hypoglycemia in a population-based cohort of children and adolescents with type 1 diabetes over the decade 2000-2009. Diabetes Care 2011;34:2379–2380

7. Miller KM, Foster NC, Beck RW, et al.; T1D Exchange Clinic Network. Current state of type 1 diabetes treatment in the U.S.: updated data from the T1D Exchange clinic registry. Diabetes Care 2015;38:971–978

8. Kowalski A. Pathway to artificial pancreas systems revisited: moving downstream. Diabetes Care 2015;38:1036–1043

9. U.S. Food and Drug Administration Center for Devices and Radiological Health. *Guidance for Industry and Food and Drug Administration Staff: The Content of Investigational Device Exemption (IDE) and Premarket Approval (PMA) Applications for Artificial Pancreas Device Systems*. Rockville, MD, U.S. Department of Health and Human Services, 2012

10. Bailey TS, Chang A, Christiansen M. Clinical accuracy of a continuous glucose monitoring system with an advanced algorithm. J Diabetes Sci Technol 2015;9:209–214

11. Damiano ER, McKeon K, El-Khatib FH, Zheng H, Nathan DM, Russell SJ. A comparative effectiveness analysis of three continuous glucose monitors: the Navigator, G4 Platinum, and Enlite. J Diabetes Sci Technol 2014;8:699–708

12. Kropff J, Bruttomesso D, Doll W, et al. Accuracy of two continuous glucose monitoring systems: a head-to-head comparison under clinical research centre and daily life conditions. Diabetes Obes Metab 2015;17:343–349

13. Peyser TA, Nakamura K, Price D, Bohnett LC, Hirsch IB, Balo A. Hypoglycemic accuracy and improved low glucose alerts of the latest Dexcom G4 Platinum continuous glucose

monitoring system. Diabetes Technol Ther 2015;17:548–554

14. Sharifi A, Varsavsky A, Ulloa J, et al. Redundancy in glucose sensing: enhanced accuracy and reliability of an electrochemical redundant sensor for continuous glucose monitoring. J Diabetes Sci Technol. 22 October 2015 [Epub ahead of print]

15. Thabit H, Leelarathna L, Wilinska ME, et al. Accuracy of continuous glucose monitoring during three closed-loop home studies under free-living conditions. Diabetes Technol Ther 2015;17:801–807

16. Wentholt IM, Vollebregt MA, Hart AA, Hoekstra JB, DeVries JH. Comparison of a needle-type and a microdialysis continuous glucose monitor in type 1 diabetic patients. Diabetes Care 2005;28:2871–2876

17. Chiang JL, Kirkman MS, Laffel LM, Peters AL; Type 1 Diabetes Sourcebook Authors. Type 1 diabetes through the life span: a position statement of the American Diabetes Association. Diabetes Care 2014;37:2034–2054

18. Bergenstal RM, Ahmann AJ, Bailey T, et al. Recommendations for standardizing glucose reporting and analysis to optimize clinical decision making in diabetes: the ambulatory glucose profile. J Diabetes Sci Technol 2013;7:562–578

19. Joubert M, Baillot-Rudoni S, Catargi B, et al.; Société Francophone du Diabète (SFD); Société Française d'Endocrinologie (SFE); EVAluation dans le Diabète des Implants ACtifs Group (EVADIAC). Indication, organization, practical implementation and interpretation guidelines for retrospective CGM recording: a French position statement. Diabetes Metab 2015;41:498–508

20. Matthaei S, Dealaiz Antuna R, Bosi E, Evans M, Geelhoed-Duijvestijn N, Joubert M. Consensus recommendations for the use of ambulatory glucose profile in clinical practice. Br J Diabetes Vasc Dis 2014;14:153–157

21. Wright E Jr, Manivannan S. Hot topics in primary care: ambulatory glucose profiling. J Fam Pract 2015;64(Suppl.):S44–S47

22. Doyle FJ 3rd, Huyett LM, Lee JB, Zisser HC, Dassau E. Closed-loop artificial pancreas systems: engineering the algorithms. Diabetes Care 2014;37:1191–1197

23. Clarke W, Kovatchev B. Statistical tools to analyze continuous glucose monitor data. Diabetes Technol Ther 2009;11(Suppl. 1):S45–S54

24. Brod M, Christensen T, Bushnell DM. The impact of non-severe hypoglycemic events on daytime function and diabetes management among adults with type 1 and type 2 diabetes. J Med Econ 2012;15:869–877

25. Brod M, Christensen T, Thomsen TL, Bushnell DM. The impact of non-severe hypoglycemic events on work productivity and diabetes management. Value Health 2011;14:665–671

26. Brod M, Pohlman B, Wolden M, Christensen T. Non-severe nocturnal hypoglycemic events: experience and impacts on patient functioning and well-being. Qual Life Res 2013;22:997–1004

27. Jensen MM, Pedersen-Bjergaard U. Self-reported frequency and impact of non-severe hypoglycemic events in insulin-treated diabetic patients in Denmark. Diabetes Management 2015;5:67–78

28. Amiel SA, Sherwin RS, Simonson DC, Tamborlane WV. Effect of intensive insulin therapy on glycemic thresholds for counterregulatory hormone release. Diabetes 1988;37: 901–907

29. Cryer PE. Mechanisms of hypoglycemia-associated autonomic failure and its component syndromes in diabetes. Diabetes 2005;54:3592–3601

30. Kropff J, Del Favero S, Place J, et al.; AP@home Consortium. 2 month evening and night closed-loop glucose control in patients with type 1 diabetes under free-living conditions: a randomised crossover trial. Lancet Diabetes Endocrinol 2015;3:939–947

31. Ly TT, Maahs DM, Rewers A, Dunger D, Oduwole A, Jones TW; International Society for Pediatric and Adolescent Diabetes. ISPAD clinical practice consensus guidelines 2014. Assessment and management of hypoglycemia in children and adolescents with diabetes. Pediatr Diabetes 2014;15(Suppl. 20): 180–192

32. Seaquist ER, Anderson J, Childs B, et al. Hypoglycemia and diabetes: a report of a workgroup of the American Diabetes Association and the Endocrine Society. Diabetes Care 2013;36: 1384–1395

33. Clarke WL, Renard E. Clinical requirements for closed-loop control systems. J Diabetes Sci Technol 2012;6:444–452

34. Del Favero S, Facchinetti A, Sparacino G, Cobelli C; AP@home Consortium. Retrofitting of continuous glucose monitoring traces allows more accurate assessment of glucose control in outpatient studies. Diabetes Technol Ther 2015; 17:355–363

35. Hovorka R, Nodale M, Haidar A, Wilinska ME. Assessing performance of closed-loop insulin delivery systems by continuous glucose monitoring: drawbacks and way forward. Diabetes Technol Ther 2013;15:4–12

36. Kollman C, Calhoun P, Lum J, Sauer W, Beck RW. Evaluation of stochastic adjustment for glucose sensor bias during closed-loop insulin delivery. Diabetes Technol Ther 2014;16:186–192

37. Beck RW, Calhoun P, Kollman C. Challenges for outpatient closed loop studies: how to assess efficacy. Diabetes Technol Ther 2013;15: 1–3

38. Kovatchev BP, Patek SD, Ortiz EA, Breton MD. Assessing sensor accuracy for non-adjunct use of continuous glucose monitoring. Diabetes Technol Ther 2015;17:177–186

39. Barnard KD, Hood KK, Weissberg-Benchell J, Aldred C, Oliver N, Laffel L. Psychosocial assessment of artificial pancreas (AP): commentary and review of existing measures and their applicability in AP research. Diabetes Technol Ther 2015;17:295–300

40. U.S. Food and Drug Administration Center for Devices and Radiological Health. *Patient Preference Information–Submission, Review in PMAs, HDE Applications, and De Novo Requests, and Inclusion in Device Labeling*. Rockville, MD, U.S. Department of Health and Human Services, 2015

41. American Diabetes Association. Introduction. In *Standards of Medical Care in Diabetes*. Diabetes Care 2011;34(Suppl. 1):S1–S2