Check for updates

Insulin dose optimization using an automated artificial intelligence-based decision support system in youths with type 1 diabetes

Revital Nimri [1], Tadej Battelino [2], Lori M. Laffel[3], Robert H. Slover[4], Desmond Schatz[5], Stuart A. Weinzimer [6], Klemen Dovc [2], Thomas Danne[7], Moshe Phillip [1,8] ✉ and NextDREAM Consortium*

Despite the increasing adoption of insulin pumps and continuous glucose monitoring devices, most people with type 1 diabetes do not achieve their glycemic goals[1]. This could be related to a lack of expertise or inadequate time for clinicians to analyze complex sensor-augmented pump data. We tested whether frequent insulin dose adjustments guided by an automated artificial intelligence-based decision support system (AI-DSS) is as effective and safe as those guided by physicians in controlling glucose levels. ADVICE4U was a six-month, multicenter, multinational, parallel, randomized controlled, non-inferiority trial in 108 participants with type 1 diabetes, aged 10–21 years and using insulin pump therapy (ClinicalTrials.gov no. NCT03003806). Participants were randomized 1:1 to receive remote insulin dose adjustment every three weeks guided by either an AI-DSS, (AI-DSS arm, $n = 54$) or by physicians (physician arm, $n = 54$). The results for the primary efficacy measure—the percentage of time spent within the target glucose range (70–180 mg dl^{-1} (3.9–10.0 mmol l^{-1}))—in the AI-DSS arm were statistically non-inferior to those in the physician arm ($50.2 \pm 11.1\%$ versus $51.6 \pm 11.3\%$, respectively, $P < 1 \times 10^{-7}$). The percentage of readings below 54 mg dl^{-1} (<3.0 mmol l^{-1}) within the AI-DSS arm was statistically non-inferior to that in the physician arm ($1.3 \pm 1.4\%$ versus $1.0 \pm 0.9\%$, respectively, $P < 0.0001$). Three severe adverse events related to diabetes (two severe hypoglycemia, one diabetic ketoacidosis) were reported in the physician arm and none in the AI-DSS arm. In conclusion, use of an automated decision support tool for optimizing insulin pump settings was non-inferior to intensive insulin titration provided by physicians from specialized academic diabetes centers.

Intensive insulin therapy is currently the standard of care for individuals with type 1 diabetes and is needed to reduce the risk of diabetes-related complications[2-4]. However, despite the increasing adoption of insulin pumps and continuous glucose monitoring (CGM) technologies[5], fewer than one-third of individuals with type 1 diabetes achieve their diabetes management goals[1,6]. Inaccurate insulin pump settings due to lack clinician knowledge or inadequate

frequency of insulin dosage adjustments could, in part, explain why glycemic control continues to worsen in these patients[7].

Frequent adjustment of insulin pump settings is particularly important in younger individuals so as to accommodate their rapidly changing insulin requirements[8]. In the landmark Diabetes Control and Complications Trial (DCCT), people with type 1 diabetes who were treated with intensive insulin therapy achieved an average 2% reduction in glycated hemoglobin by attending weekly in-person clinic visits until glycemic goals were met, followed by monthly visits and weekly telephone contacts[2,9]. These findings and others support that regular and frequent insulin dose adjustments are needed to improve glycemic control in a broad population of people with either type 1 diabetes or type 2 diabetes[10-12]. Nevertheless, current guidelines recommend that children, adolescents and young adult individuals with type 1 diabetes attend clinical visits every 3–4 months[13,14]. Current standard of care guidelines are thus inadequate. Additionally, patient adherence to this recommendation is often poor[15]. Obtaining pump and continuous glucose data then analyzing the data are also time-consuming tasks for physicians, presenting an additional barrier to the provision of frequent insulin pump adjustments[16].

Given the mounting shortage of endocrinologists, especially in rural areas[17], the current paradigm of clinical care lacks the ability to provide the timely insulin adjustments needed to improve the management of type 1 diabetes. Digital decision support systems can address these issues by facilitating timely and more frequent insulin dose adjustments, either in person or remotely, and thus improve glycemic control for individuals with diabetes.

We evaluated an automated artificial intelligence-based decision support system (AI-DSS; The DreaMed Advisor Pro; DreaMed Diabetes Ltd., Petah Tikva, Israel), which provides insulin therapy adjustment recommendations and tips for diabetes management to healthcare professionals managing individuals with type 1 diabetes using an insulin pump and CGM. Preliminary studies provided initial indications of the safety and clinical validity of using the AI-DSS device[18]. A survey of 26 physicians with experience in insulin pump dosing titration showed that the level of agreement among physicians

[1]The Jesse Z and Sara Lea Shafer Institute for Endocrinology and Diabetes, National Center for Childhood Diabetes, Schneider Children's Medical Center of Israel, Petah Tikva, Israel. [2]Department of Endocrinology, Diabetes and Metabolic Diseases, UMC-University Children's Hospital Ljubljana, and Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia. [3]Joslin Diabetes Center, One Joslin Place, Harvard Medical School, Boston, MA, USA. [4]Barbara Davis Center for Childhood Diabetes, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. [5]Department of Pediatrics, College of Medicine, University of Florida, Gainesville, FL, USA. [6]Pediatric Endocrinology & Diabetes, Yale School of Medicine, New Haven, CT, USA. [7]Diabetes Center for Children and Adolescents, Children's Hospital AUF DER BULT, Hannover, Germany. [8]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: mosheph@tauex.tau.ac.il
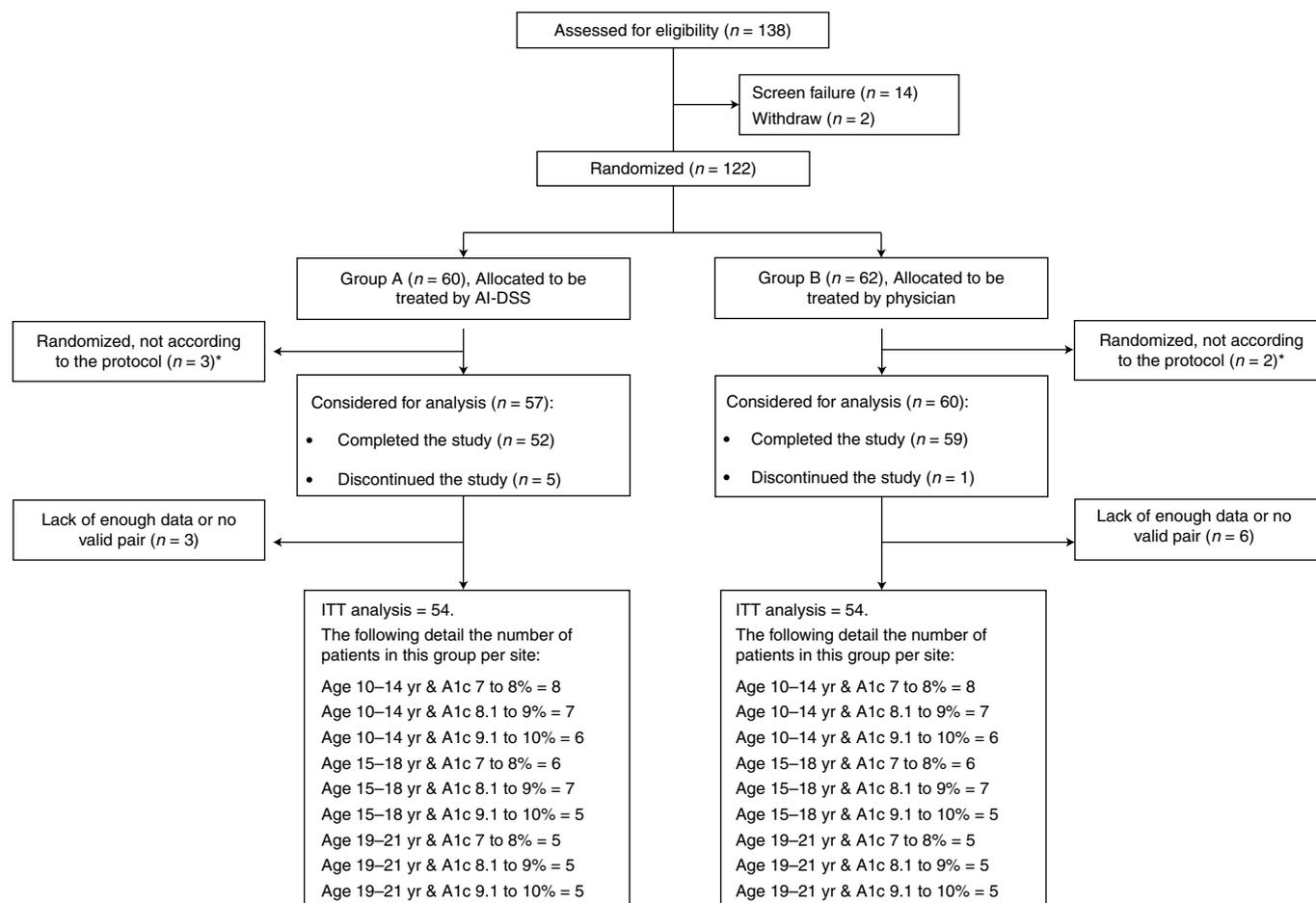
Fig. 1 | Study flow chart. Asterisks indicate not randomized according to designed hemoglobin A1c strata at the site. ITT, intention to treat.

regarding the direction of insulin recommendations (increase, decrease or no change) for 15 patients with type 1 diabetes was only around 41–46% for each of the pump parameters (basal insulin infusion rate, the insulin-to-carbohydrate ratio (CR) and the correction factor (CF)) and the level of disagreement was around 10–12% (ref. [18]). In the same study, similar levels of agreement were found between the physicians and the AI-DSS. We hypothesized that frequent optimization of insulin pump therapy based on CGM readings using the AI-DSS would result in non-inferior glycemic control compared with dose adjustments performed by physicians from specialized academic diabetes centers.

Results

From 20 November 2017 to 25 July 2019, 138 individuals were assessed for eligibility; 122 were randomized and 108 participants (AI-DSS, $n = 54$; physician, $n = 54$) were included in the intention-to-treat analysis (Fig. 1). The demographic and clinical characteristics of the randomized participants were similar between the two arms (Table 1). Sixty participants were randomized at US sites and 62 at European and Israeli sites. The disposition of randomized participants is presented in Fig. 1. The average (s.d.) number of CGM-valid days for analysis was 137 ± 35 days per participant in the 'physician' arm and 133 ± 32 days in the AI-DSS ('advisor') arm ($P = 0.25$, two-tailed paired t-test).

The percentage of time spent within target glucose range in the AI-DSS arm was statistically non-inferior to the physician arm ($50.2 \pm 11.1\%$ versus $51.6 \pm 11.3\%$, respectively) (95% lower confidence bound CI: -3.3%, $P < 1 \times 10^{-7}$). The percentage of

readings below $54\,\mathrm{mg\,dl^{-1}}$ was statistically non-inferior between the AI-DSS and physician arms ($1.3 \pm 1.4\%$ and $1.0 \pm 0.9\%$, respectively. $P < 0.0001$).

A total of 20 diabetes-related study adverse events (AEs) (AI-DSS arm, $n = 8$; physician arm, $n = 12$) were reported. Among these, there were three severe AEs (two severe hypoglycemia, one diabetic ketoacidosis), which were reported in the physician arm. All AEs are presented in Table 2.

Both arms showed statistically significant reductions in mean glycated hemoglobin level from baseline to mid study (week 12; Fig. 2). A statistically significant reduction of 0.32% in mean glycated hemoglobin level from baseline to end of study (week 24) was observed in the AI-DSS arm (two-sided 95% CI of -0.55% to -0.08%, $P = 0.008$). The reduction from baseline to week 24 in the physician arm was 0.19% (two-sided 95% CI of -0.49% to 0.11%, $P = 0.22$; Fig. 2). However, the change in glycated hemoglobin level from baseline to the end of the study when comparing the two arms was found to be statistically not different ($P = 0.49$).

No between-arm differences were observed in the percentage of time above and below glucose target ranges (Extended Data Fig. 1). Mean total daily insulin and daily basal insulin doses were not statistically different. Daily bolus insulin doses among participants in the AI-DSS arm were higher than in the physician arm (29.6 ± 9.7 and 26.6 ± 8.3 units, respectively, $P = 0.03$).

The per protocol outcomes, which included 30 participants in each arm, are presented in Extended Data Figs. 2 and 3.

Eight of the 13 AI-DSS physicians completed the satisfaction questionnaire at week 12 and all completed the questionnaire at

Table 1 | Randomized participant characteristics at baseline

Characteristic[a]	AI-DSS arm (n = 60)	Physician arm (n = 62)
Gender (F/M)	32/28	32/30
Age (yr)	15.5 ± 3.0	15.8 ± 3.0
Weight (kg)	61.7 ± 13.8	63.4 ± 13.1
Height (cm)	164.3 ± 11.0	167.0 ± 11.0
BMI[b]	22.6 ± 3.4	22.5 ± 3.1
BMI-SDS[c]	0.5 ± 0.9	0.7 ± 0.7
Glycated hemoglobin (%)	8.4 ± 0.8	8.4 ± 0.8
Glycated hemoglobin (mmol mol^{-1})	68.4 ± 8.5	68.0 ± 8.8
Total daily insulin (U kg^{-1} d^{-1})[d]	0.9 ± 0.2	0.8 ± 0.2
Diabetes duration (yr)	6.6 ± 4.1	7.7 ± 4.2
Pump-therapy duration (yr)	4.9 ± 3.8	5.4 ± 3.7
Pump brand used:		
Medtronic VEO, 530G, 640G, 630G	31	33
Omnipod	29	28
Not recorded	0	1
Sensor use duration (yr)	1.9 ± 1.9	2.6 ± 2.5
Sensor brand used before study start:		
DexCom G4/G5	43	40
Medtronic Enlite	12	15
FreeStyle Libre	2	5
FreeStyle Navigator II	2	1
None	1	1

[a]Values presented as mean ± s.d. [b]The body mass index (BMI) is the weight in kilograms divided by the square of the height in meters. [c]The BMI-SDS was calculated in 43 pediatric participants in the AI-DSS arm and 44 participants in the physician arm. [d]Baseline insulin information was recorded from 56 participants in the AI-DSS arm and 60 participants in the physician arm.

Table 2 | Adverse events

	AI-DSS arm (N = 60)	Physician arm (N = 62)
No. of severe hypoglycemic events[a]	0	2
No. of severe hyperglycemic event (diabetic ketoacidosis)	0	1
No. of severe AEs unrelated to diabetes[b]	2	1
Significant hyperglycemia[c] (due to pump malfunction)	2 (1)	8 (4)
Ketonuria	0	2
Significant hypoglycemia[d]	3	2
No. of device-related AEs		
Sensor-related contact allergic	1	0
Insulin pump site infection	0	4
No. of AEs not related to study interventions (sum)	44	55
Ear, throat and respiratory infections[e]	21	28
Gastrointestinal infections and inflammatory conditions[f]	8	8
Bone and muscle and joint injury or pain	3	7
Allergic conditions	3	0
Urinary infections	2	0
Conjunctivitis	0	2
Skin and subcutaneous tissue disorders	2	4
Neurologic (syncope/headache)	2	0
Other[g]	3	6

[a]Severe hypoglycemia defined as a clinical episode of hypoglycemia, resulting in seizure or coma, requiring intravenous glucose or glucagon administration or any hypoglycemia that requires assistance from another person or intervention by hospital. [b]Severe AEs that were not related to study interventions, including a surgery of pilonidal abscess and a surgery of tonsillectomy and adenoidectomy in two patients and hospitalization due to supraventricular tachycardia in one patient. [c]Significant hyperglycemia reported that required assistance to treat hyperglycemia and significant urine ketones levels above +3. [d]Significant hypoglycemia reported at events with prolonged hypoglycemia or glucose level below 33 mg dl^{-1} and not meeting the criteria for severe hypoglycemia. [e]Upper respiratory tract infection, nasopharyngitis, laryngitis, sinusitis, common cold/nasal congestion, bronchitis, otitis, ear pain. [f]Abdominal pain, dyspepsia, gastroenteritis [g]Tachycardia, dehydration, fatigue, viral infection unspecified.

week 24. At study end, the majority (n = 11, 85%) stated they were interested in continuing to use the AI-DSS as part of their routine practice. Responses to the Healthcare Professional Post-Intervention Survey are presented in Extended Data Fig. 4.

Overall, 74.9% of the planned insulin dose titration visits were performed in the AI-DSS arm and 87.7% in the physician arm. The main reason for not providing a dose titration was due to insufficient or missing glucose and/or insulin data. Descriptions of the numbers and technical issues related to missing insulin dose visits recommendations in both study arms are presented in Extended Data Fig. 5.

Override information was available for 50 of 54 participants in the AI-DSS arm (four participants did not receive any recommendations throughout the study; see Extended Data Fig. 5 for further details). A total of 12,864 comparison time points were available for each pump setting parameter. The percentage of overrides was 3.2% for the basal plan, 1.7% for the CF and 0.9% for the CR. The overall percentage of override was 1.9% for the combined pump settings parameters. The absolute relative differences in the magnitude of change made were 10 ± 6.3%, 6.2 ± 7.3% and 15.5 ± 4.7% for basal, CF and CR, respectively.

Discussion

In this randomized, controlled, single-masked, non-inferiority trial, we found that the use of the AI-DSS for optimization of insulin pump settings was non-inferior in efficacy and safety to intensified clinical care provided by trained physicians. The AI-DSS was

found to be safe and the percentage of readings in the hypoglycemia range below 54 mg dl^{-1} within the AI-DSS arm was statistically non-inferior to that in the physician arm.

Frequent insulin dose adjustments by healthcare professionals between in-person visits using downloaded data from devices have been shown to improve glycemic control[12,19]. The current study was not designed to evaluate the effect of dose titration; however, a significant reduction in glycated hemoglobin level was observed in both arms after 12 weeks, with a significant reduction using the AI-DSS at the end of the study.

There was a high level of satisfaction among the physicians who used the AI-DSS and who completed the Healthcare Professional Post-Intervention Survey. However, there were cases where physicians assigned to the AI-DSS arm decided to change the recommendations given by the AI-DSS. Because the percentage of these overrides was relatively small, their influence on the outcomes is likely to be negligible. That some physicians override the AI-DSS recommendations is not surprising. Insulin dose adjustments are not only influenced by science but also by different subjective attitudes—individual expert clinicians have been shown previously to provide different insulin dosing adjustments to the same given patient data[18].
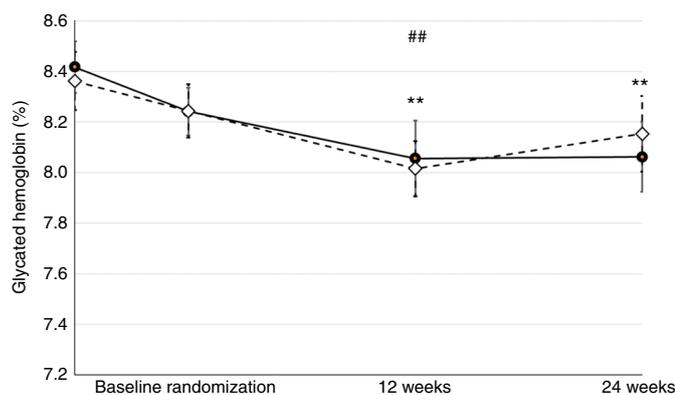
Fig. 2 | Glycated hemoglobin levels at baseline, randomization, 12 weeks and 24 weeks for the intent to treat ($n = 53$ per arm) cohort. Values are expressed as means (markers) ± standard error (error bars). Filled circles mark the AI-DSS arm and open diamonds the physician arm. The results of two-sided, paired t-test analyses are marked by asterisks to indicate a statistically significant difference between baseline and a specific time point within the AI-DSS arm and by hash symbols to indicate a statistically significant difference between baseline and a specific time point within the physician arm (where shown, these indicate $P < 0.01$).

A key strength of our study was the use of a non-inferiority design to evaluate a new AI-DSS by comparison with a known treatment approach provided by physicians from academic centers who are experienced in the use of diabetes technologies. Another strength of our study was the randomization of study physicians, which helped to ensure equal levels of expertise in diabetes management between the study arms. Furthermore, participants were masked to their randomization arm, eliminating bias regarding the source of the given insulin recommendations (automated versus physician). Additionally, stratification of participants according to age and baseline glycated hemoglobin level facilitated a more robust comparison of closely matched study arms. Finally, use of a multicenter, multinational design further supports the generalizability of the study findings.

A potential limitation is the effort involved in frequent uploading and downloading of data as well as the minimum CGM data required for the scheduled, obligatory, insulin titrations. Nonetheless, the relatively low number of study withdrawals probably indicates that the benefits of frequent insulin titration and the ease with which the recommendations were communicated to the participants could outweigh the burden of data sharing. Moreover, several technological advancements already exist that potentially mitigate this limitation (for example, pre-calibrated CGM and connected devices). In addition, the remote nature of the data sharing and subsequent dose adjustments through online or mobile applications were probably less disruptive than frequent in-person clinic visits, which often require absences from school or work. Real-world studies are needed to fully explore the persistence of physicians and individuals with diabetes in using the AI-DSS in daily diabetes management. The study is also limited by the fact that it included a selected population of young people with type 1 diabetes who were already receiving their diabetes care at academic tertiary care institutions and who were willing to participate in the study by providing written informed consent and satisfying the inclusion and exclusion criteria. These individuals might not generally represent young people with type 1 diabetes, which could limit the ability to generalize these data to broader populations. An additional limitation of the study is that the AI-DSS is suitable for a specific population of people with diabetes who have the ability to upload the device's diabetes-related

data and who have regular access to the internet. The challenge of manual data uploading and downloading could be eliminated by real-time data transmission in the near future; however, an internet connection will still be required.

Given the increasing prevalence of diabetes[20,21] and the mounting shortage of specialized endocrinologists[22], the responsibility to manage individuals with diabetes may be increasingly taken on by primary care physicians. Because many of these physicians are inexperienced in insulin therapy and the use of CGM data, innovative and readily available tools will probably be welcome additions to support clinical decision making in the primary care setting. Decision support systems have the potential to deliver this assistance and to elevate the quality of diabetes care by creating a virtual expert diabetes clinic that facilitates more frequent insulin adjustments between clinic visits. These systems can be used with telemedicine approaches to deliver expert knowledge, reduce disease burden for patients and alleviate burdens on caregivers. They can also lower the costs associated with clinical visits (for example, lost work and/or school days) and overcome missed visits[15], particularly among people who live in rural areas[17] or during other circumstances in which access to face-to-face visits with physicians is limited. Moreover, the AI-DSS provides the ability to standardize care across health systems, ensuring that all people with type 1 diabetes would receive quality care, even if they live in remote areas or are managed by a care team with little experience with CGM interpretation or insulin pump therapy.

Our study shows that optimizing insulin pump settings with an AI-DSS at three-week intervals for six months achieved statistically similar levels of glucose control to those achieved by physicians with diabetes expertise. The AI-DSS used in our study provides opportunities for a new modality for intensive insulin management that can offer the frequently needed insulin adjustments necessary to improve glucose control in young people with type 1 diabetes.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-020-1045-7.

References

1. Miller, K. M. et al. Current state of type 1 diabetes treatment in the US: updated data from the T1D exchange clinic registry. *Diabetes Care* **38**, 971–978 (2015).
2. 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.* **329**, 977–986 (1993).
3. Nathan, D. M. et al. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N. Engl. J. Med.* **353**, 2643–2653 (2005).
4. American Diabetes Association. Standards of medical care in diabetes—2019. *Diabetes Care* **42**, S1–S193 (2019).
5. Foster N. C. et al. Marked increases in CGM use has not prevented increases in HbA1c levels in participants in the T1D exchange (T1DX) clinic network. *Diabetes* **67** (2018); https://doi.org/10.2337/db18-1689-P
6. Edelman, S. V., Argento, N. B., Pettus, J. & Hirsch, I. B. Clinical implications of real-time and intermittently scanned continuous glucose monitoring. *Diabetes Care* **41**, 2265–2274 (2018).
7. Zisser, H. et al. Clinical performance of three bolus calculators in subjects with type 1 diabetes mellitus: a head-to-head-to-head comparison. *Diabetes Technol. Ther.* **12**, 955–961 (2010).
8. Petitti, D. B. et al. Glycemic control in youth with diabetes: the SEARCH for Diabetes in Youth Study. *J. Pediatr.* **155**, 668–672 (2009).
9. The DCCT Research Group. Diabetes Control and Complications Trial (DCCT): results of feasibility study. *Diabetes Care* **10**, 1–19 (1987).

10. Davidson, M. B. How our current medical care system fails people with diabetes: lack of timely, appropriate clinical decisions. *Diabetes Care* **32**, 370–372 (2009).

11. Bergenstal, R. M. et al. Automated insulin dosing guidance to optimise insulin management in patients with type 2 diabetes: a multicentre, randomised controlled trial. *Lancet* **393**, 1138–1148 (2019).

12. Shalitin, S. et al. Using the Internet-based upload blood glucose monitoring and therapy management system in patients with type 1 diabetes. *Acta Diabetol.* **51**, 247–256 (2014).

13. Pihoker, C. et al. ISPAD clinical practice consensus guidelines 2018: the delivery of ambulatory diabetes care to children and adolescents with diabetes. *Pediatr. Diabetes* **19**, 84–104 (2018).

14. Chiang, J. L. et al. Type 1 diabetes in children and adolescents: a position statement by the American Diabetes Association. *Diabetes Care* **41**, 2026–2044 (2018).

15. Markowitz, J. T., Volkening, L. K. & Laffel, L. M. B. Care utilization in a pediatric diabetes clinic: cancellations, parental attendance and mental health appointments. *J. Pediatr.* **164**, 1384–1389 (2014).

16. Comellas, M. J. et al. Evaluation of a new digital automated glycemic pattern detection tool. *Diabetes Technol. Ther.* **19**, 633–640 (2017).

17. Lu, H. et al. Population-based geographic access to endocrinologists in the United States, 2012. *BMC Health Serv. Res.* **15**, 541 (2015).

18. Nimri, R. et al. Adjusting insulin doses in patients with type 1 diabetes who use insulin pump and continuous glucose monitoring: variations among countries and physicians. *Diabetes Obes. Metab.* **20**, 2458–2466 (2018).

19. Beck, R. W. Downloading diabetes device data: empowering patients to download at home to achieve better outcomes. *Diabetes Technol. Ther.* **17**, 536–537 (2015).

20. *IDF Diabetes Atlas*, 8th edn (International Diabetes Federation, accessed 3 December 2018); https://www.idf.org/e-library/epidemiology-research/diabetes-atlas/134-idf-diabetes-atlas-8th-edition.html

21. Patterson, C. C. et al. Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989–2013: a multicentre prospective registration study. *Diabetologia* **62**, 408–417 (2019).

22. Vigersky, R. A. et al. The clinical endocrinology workforce: current status and future projections of supply and demand. *J. Clin. Endocrinol. Metab.* **99**, 3112–3121 (2014).

NextDREAM Consortium

Moshe Phillip[1], Revital Nimri[1], Shlomit Shalitin[1], Rachel Bello[1], Michal Nevo-Shenker[1], Naama Fisch-Shvalb[1], Galit Shiovitch-Mantzuri[1], Orit Choresh[1], Irit Drutz[1], Yehiel Nava[1], Alona Hemo[1], Orna Hermon[1], Rachel Nave[1], Tadej Battelino[2], Klemen Dovc[2], Natasa Bratina[2], Darja Smigoe-Schweiger[2], Brigita Mali[2], Ana Gianini[2], Urska Sever[2], Barbara Murn Berkopec[2], Lori M. Laffel[3], Michelle Katz[3], Elvira Isganaitis[3], Sanjeev Mehta[3], Heidi Quinn[3], Nisha Naik[3], Zijing Guo[3], Lisa Volkening[3], Robert H. Slover[4], Gregory Forlenza[4], R. Paul Wadwa[4], G. Todd Alonso[4], Laurel Messer[4], Lindsey Towers[4], Katie Thivener[4], Cari Berget[4], Samantha Lange[4], Emily Jost[4], Maria Rossick-Solis[4], Desmond Schatz[5], Michael Haller[5], Paul Hiers[5], Laura Jacobsen[5], Madison Smith[5], Anastasia O'Neill[5], Jennifer Hosford[5], Alexis Perry[5], Stuart A. Weinzimer[6], Eda Cengiz[6], Jennifer Sherr[6], Kathryn Gibbons[6], Lori Carria[6], Melinda Zgorski[6], Thomas Danne[7], Torben Biester[7], Olga Kordonouri[7], Thekla von dem Berge[7], Sarah Biester[7] and Kerstin Remus[7]

Methods

Trial design and oversight. The ADVICE4U study was designed to evaluate whether frequent insulin dose titration using a new treatment modality—an automated decision support system (AI-DSS arm)—would be comparable in efficacy to current best practice, that is, titration by physicians from specialized academic diabetes centers (physician arm). Accordingly, a non-inferiority design was preferred over a superiority design[23]. The ADVICE4U study was a six-month, multicenter, multinational, single-blind, parallel (two-arm), randomized controlled, prospective, non-inferiority trial that evaluated the equivalence of frequent insulin dose adjustments at three-week intervals on the percentage of time in the target glucose range (70–180 mg dl⁻¹, 3.9 to 10 mmol l⁻¹) made using a decision support system (AI-DSS arm) compared with adjustments made by physicians from academic diabetes centers including certified endocrinologists, diabetologists and fellows in pediatric endocrinology (physician arm). The trial was conducted at seven clinical sites; four in the USA (University of Florida, Joslin Diabetes Center, Barbara Davis Center, Yale University School of Medicine), two in Europe (University Children's Hospital Ljubljana Slovenia, Children's Hospital AUF DER BULT, Germany) and one in Israel (Schneider Children's Medical Center of Israel). A complete list of centers and center staff is provided in the Acknowledgments section.

The trial was conducted in compliance with the International Conference on Harmonisation/Good Clinical Practice (ICH/GCP) applicable regulatory requirements, in accordance with International Organization for Standardization (ISO) 14155:2012, the Declaration of Helsinki and Code of Federal Regulations title 21, parts 50, 54, 56 and 812. The study protocol was approved by the local institutional review boards at each participating center as follows: Joslin Diabetes Center, Committee of Human Studies; Barbara Davis Center for Childhood Diabetes, Colorado Multiple Institutional Review Board; Yale University School of Medicine, Yale University Human Investigation Committee; Diabetes Institute University of Florida, Institutional Review Board University of Florida; Schneider Children's Medical Center of Israel, Institutional Review Board Rabin Medical Center; University Children's Hospital Ljubljana Slovenia, national regulatory authorities—Republic of Slovenia National Medical Ethics Committee (JAZMP); Children's Hospital AUF DER BULT, national regulatory authority—Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Protocol version 1.5 (available in the Supplementary Information) was used for all sites. The German site was the last site to join the study as the IRB approval was delayed and used protocol version 1.6. According to a BfArM request, two changes were made to protocol version 1.5: the exclusion criteria included hypoglycemia unawareness and an additional term for premature study discontinuation: occurrence of three device/advice related SAEs during the study conduct (SAE, serious adverse event). All participants (or parents/legal guardians) provided written informed consent, and assent according to local requirements.

An external contract research organization coordinated the trial, performed site monitoring and ensured data collection and management. Safety aspects of the study were monitored by an independent data and safety monitoring board.

Participants. Children, adolescents and young adults who had had type 1 diabetes for at least one year and were treated with insulin pump therapy were eligible for participation in the trial. Inclusion criteria for participation were as follows: subjects with type 1 diabetes for at least a year since diagnosis; ages 10 to 21 years; glycated hemoglobin level from 7.0% (53 mmol mol⁻¹) to 10.0% (86 mmol mol⁻¹); use of insulin pump therapy for at least four months (OmniPod insulin pump (Insulet Corp.) or Medtronic pump: MiniMed 530G System (MMT-551, MMT-751), MiniMed Paradigm REAL-Time Revel (MMT-523, MMT-723, MMT-523K, MMT-723K) and MiniMed Paradigm (MMT-515, MMT-715, MMT-522, MMT-722, MMT-522K, MMT-722K), MiniMed Veo Insulin Pump (MMT-754, MMT-554) or any other Medtronic pump that is compatible with the diabetes management system (DMS) app used in this study); BMI-SDS below the 97th percentile for age; willingness to follow study instructions (that is, to measure at least two capillary blood glucose readings per day as needed for CGM calibration and use the bolus calculator feature of the pump); patients/parents were required to have a minimum level of computer skills and understanding of navigating the internet; willingness to use a Dexcom CGM device throughout the study duration; patients/parents have a smartphone (Apple or Android and Windows).

Exclusion criteria were an episode of diabetic ketoacidosis within the month before study entry; any significant diseases/conditions including psychiatric disorders and substance abuse that in the opinion of the investigator were likely to affect the subject's ability to complete the study or compromise patient safety; current participation in any other interventional study; known or suspected allergy to trial products such as adhesives, tapes and needles; female subject who is pregnant or lactating or planning to become pregnant within the planned study duration; severe hypoglycemia within six months prior to enrollment (as defined by the ADA and Endocrine Society as follows: severe hypoglycemia is an event requiring assistance of another person (due to change in mental status) to actively administer carbohydrates, glucagon or take other corrective actions); current use of medications that are used to lower blood glucose, such as pramlintide, metformin and GLP-1 analogs, or beta blockers, glucocorticoids and other medications, which in the judgment of the investigator would be a contraindication to participation

in the study (anticoagulant therapy such as plavix, low-molecular-weight heparin, coumadin, immunosuppressant therapy); subjects who have a relevant severe organ disorder (diabetic nephropathy, diabetic retinopathy, diabetic foot syndrome) or any secondary disease or complication of diabetes mellitus (for example, subjects who have unstable or rapidly progressive renal disease or are receiving dialysis; subjects who have active proliferative retinopathy or active gastroparesis); subjects who are suffering from an eating disorder.

Trial devices and interventions. The AI-DSS utilizes data from glucose monitoring (CGM readings and capillary blood glucose measurements), as well as insulin doses and carbohydrate intake obtained from the insulin pump data and the pump bolus calculator over at least 12 days of routine diabetes care. The AI-DSS uses artificial intelligence that detects and analyzes glucose patterns and insulin dosing events in a similar approach to that used by a healthcare provider based on expert knowledge, recommendations and data acquired from various clinical studies. Its recommendations include exact adjustments for insulin pump settings (basal rates, CRs and CFs) and it provides personalized diabetes management tips related to the way insulin is delivered (for example, missed boluses, timing of pre-meal bolus, overtreatment of hypoglycemia and so on), highlighting patterns related to an individual's insulin dosing and delivery behaviors.

The AI-DSS comprises a cloud-based server and an algorithm module. The cloud-based server is a secured database and computing platform that hosts the data inputs from the DMS (Glooko Platform—mobile app and online software; Glooko Inc.). It is used to retrieve anonymous data from the DMS, temporarily store raw data, transmit these data to the algorithm for processing and result generation, and store this to the AI-DSS results. The result is then sent to the DMS, which generates a report for the healthcare provider, who can edit, approve and share the recommendation with the patient (see Extended Data Fig. 6 for the data flow).

The algorithm is designed to provide a comprehensive analysis of an individual's diabetes data consisting of glucose levels, insulin delivery history and meal consumption, as reported through the insulin pump's bolus calculator, to recommend adjustments to the patient-specific insulin pump settings (CR, CF and basal plans) as well as suggestions for personalized diabetes management tips. The system pulls the most recent 21 days of data from the DMS from the last date of data upload. Of the data pulled, AI-DSS requires at least 12 valid days to produce recommendations. A valid day should include at least 67% of CGM sensor readings according to the sensor's sampling rate and at least one basal and one bolus record. In addition, it requires at least three records from the bolus calculator and a valid recent insulin pump settings plan (that is, the basal plan is set and the CR, CF, glucose targets and active insulin are set inside the pump bolus calculator). The AI-DSS uses the raw data input to detect dosing decision events by the patient using both the insulin pump and CGM glucose meter to characterize each insulin dosing event. In parallel, it detects patterns of hypoglycemia, euglycemia and hyperglycemia. The event-based analysis is then integrated with the patient's glycemic patterns to produce a personalized recommendation. The recommended changes to the patient's pump settings are limited to ensure safety: basal rate changes are limited to ±20% of the current hourly rate and no more than 50–150% of the hourly average basal rate calculated from the patient total daily insulin dose. The CR and CF changes are limited to ±30% of the current insulin pump settings.

The AI-DSS recommendations are created in a few seconds after the participant uploads the device's data. The recommendation is sent to the physician for review and can be shared with the patient. An example of an AI-DSS recommendation report for the healthcare provider is presented in Extended Data Fig. 7.

Once the healthcare provider has approved a recommendation, it is shared with the patient through the AI-DSS section in the DMS app. The patients will receive a push notification to their phone as a reminder that they have a new recommendation. The patient can review the current recommendation (Extended Data Fig. 8) and agree or disagree with it. In the case of disagreement, the patient can document the reason and send feedback to the healthcare provider.

The DMS with integrated algorithm or physician recommendation report was used in both study arms for downloading the pump and CGM data, for a visual display of current insulin pump settings and for recording recommended insulin pump dose changes according to the designated study arm (allowing edits as needed), as well as uploading of the recommendations to share with participants/families. The DreaMed Advisor Pro algorithm version 01.07.xx was used during the study.

This was an investigator-initiated study and the AI-DSS was provided by DreaMed Diabetes Ltd for the study. Because the investigators did not have access to the device's internal database, which is essential for error analysis, no such analysis was performed as part of the study. In addition, the AI-DSS software has been tested extensively as part of an FDA regulatory submission. DreaMed Diabetes Ltd reported no need to use the data collected in this study for further error analysis.

Procedures. The trial consisted of a three-week run-in period in which participants continued their regular treatment and used the provided study CGM

device. At the end of the run-in period, participants were randomized either to the AI-DSS or physician arms. Three weeks after randomization and every three weeks thereafter, all participants uploaded their insulin pump and CGM data using the DMS. Participants then received insulin pump settings recommendations either from the physician or AI-DSS according to their randomization, for a 24-week period. The number of scheduled contacts with study staff was identical for both study arms. The study design scheme and visit schedule are presented in Extended Data Figs. 9 and 10, respectively.

At the screening visit, baseline characteristics were collected, and all participants received a DexCom G5 CGM system (Dexcom) to use throughout the study. Participants continued to use their insulin pump (Medtronic MiniMed 530G, 640G/630G, Veo or Paradigm Realtime-Revel (all Medtronic Diabetes) or an OmniPod (Insulet Corporation)).

After a three-week run-in period, eligible participants were randomly assigned in a 1:1 ratio to either the AI-DSS or physician arms, using computer-generated permuted block randomization stratified by site, age (10–14 years, 15–17 years, 18–21 years) and glycated hemoglobin level at screening (7 to <8% (53 to <64 mmol mol^{-1}), 8 to <9% (64 to <75 mmol mol^{-1}) and 9–10% (75–86 mmol mol^{-1})) to achieve an equal distribution of participants in the two study arms. In this participants stratification there were thus nine participant strata. As the sample size calculation showed that 56 pairs of participants would give the study sufficient power and the study included seven centers, each site was assigned eight of the nine age-glycated hemoglobin strata and asked to recruit one pair of participants in each stratum (for example, each pair had one participant designated to the AI-DSS arm and the other to the physician arm). The decision as to which eight strata to allot to each center was done at random, with the constraint of keeping the number of pairs in each stratum as equal as possible. The purpose of using this kind of randomization was to ensure the inclusion of a wide range of ages with different levels of glycemic control in an equal fashion. On 3 September, 2018, open baskets for randomization were open for competitive recruitment. In addition, each clinical site included four physicians (except for one site that had two physicians) who were split into pairs, with each pair possessing similar levels of training and experience. Each pair of physicians was randomly assigned to review pump settings either by themselves (physician arm) or by using the AI-DSS (AI-DSS arm) for the study. It is important to emphasize that none of the physicians who actively participated in the study provided expert advice and/or guidance that informed the AI-DSS design. Additional information is presented in the study protocol.

Three weeks after randomization and every three weeks thereafter, for a total of seven times, participants uploaded their insulin pump and CGM data using the DMS. Extended Data Fig. 6 illustrates the flow of the data in the study. The DMS pulled the last 21 days of insulin pump and CGM data. The data were transferred to the AI-DSS; if the source of the data was a participant in the physician arm, a report form was created without insulin dosing recommendations for the physician to fill in electronically with their dosing recommendations. If the data came from a participant in the AI-DSS arm, the system verified that the data requirements included at least 12 valid days before producing recommendations. A valid day for the AI-DSS system is defined as a day with at least 67% of CGM data and at least one basal record and one bolus record. If the minimal data requirements were met, the AI-DSS produced insulin pump settings recommendations, which were then shown to the physician in the report form. Thus, at this point, in both arms, a report including the insulin pump and CGM data was available in the DMS for the designated physician to review. The AI-DSS arm physicians also received the automated insulin recommendations for review (for approval or revision by a designated physician). Physicians in the physician arm analyzed participant data and determined insulin pump adjustments based on their clinical judgment. For both arms, the recommended adjustments were then shared with the participant and/or family via the DMS, email and/or in a follow-up phone call in a similar manner, maintaining the masking of group assignment to participants. An example of the recommendations report as viewed in the DMS web platform for the physicians is presented in Extended Data Fig. 7 and the recommendations report for the participants as viewed through their smartphone is shown in Extended Data Fig. 8. Participants were instructed to enter the recommended pump settings and download the pump data within 24 h to confirm that the new settings had been correctly entered.

Glycated hemoglobin was measured four times throughout the study (at baseline, randomization, week 12 and week 24) using a point-of care standardized capillary glycated hemoglobin measurement (DCA 2000, DCA 2000+ or DCA Vantage, utilizing the same brand of reagents). Physicians participating in the AI-DSS arm were surveyed for their satisfaction at week 12 and week 24.

Outcomes. The primary efficacy outcome was the non-inferiority for percentage of time that the glucose level, as measured by the continuous glucose monitor, was in the target range of 70–180 mg dl^{-1} (3.9–10.0 mmol l^{-1}) over the active treatment period between the two arms. The primary safety outcome was the percentage of time that the glucose level was less than 54 mg dl^{-1} (3.0 mmol l^{-1}). Secondary outcomes were changes in glycated hemoglobin level from baseline to the end of study and AEs and SAEs. The exploratory glycemic measures were mean glucose, glycemic variability as measured by standard deviation (s.d.) and coefficient of variation (%CV), percentage of time that glucose level readings were <50 mg dl^{-1}

(<2.8 mmol l^{-1}), <70 mg dl^{-1} (<3.9 mmol l^{-1}), >180 mg dl^{-1} (>10 mmol l^{-1}) and >240 mg dl^{-1} (13.3 mmol l^{-1}), insulin doses, as well as overrides of AI-DSS recommendations done by the physicians in the AI-DSS arm. Device satisfaction was evaluated using a 50-item investigator-developed questionnaire (Healthcare Professional Post-Intervention Survey) completed by the physicians who used the AI-DSS during the study (at 12 and 24 weeks). The psychosocial questionnaire completed by the healthcare providers during the study had good face validity, but its psychometric properties are unknown as it was developed in collaboration with participants, specifically for this study. Future research aims to validate the questionnaire.

Statistical analysis. We hypothesized that use of the AI-DSS would not be inferior to physician-guided recommendations in percentage of sensor readings within the target range of 70–180 mg dl^{-1} (3.9–10 mmol l^{-1}). For the sample size calculation, data from the JDRF CGM randomized controlled study[24] and the REPLACE-BG study[25] were used to estimate the properties of the primary endpoint and to define the non-inferiority limit. The sample size was calculated using WINPEPI[26] for non-inferiority analysis, with a margin of 7.5% between arms in a parallel study design with power of 90%, a one-sided significance level of 5% and s.d. of 13% with a correlation of 0.48 between the baseline and outcome time in range over the 24-week treatment period. The resulting analysis showed that 104 subjects were needed for the study. The sample size was set at 112 participants (16 patients at each site), given an estimated 5–10% dropout rate, with randomization in a 1:1 ratio between the two study arms.

The analyses complied with the intention-to-treat principle, including participants who were recruited according to the randomization rules and those pairs of participants for which follow-up data were available for both members of the pair. All principal analyses were based on paired comparisons, with pairs formed from participants within the same center. This methodology was used to eliminate inter-center differences. The relationship between the percentage of time in range and the treatment arm was examined by analysis of covariance (ANCOVA), adjusting for baseline percentage of time in range and number of valid days (a valid day is defined as a day with at least 67% of CGM values) used in the analysis. A one-sided hypothesis test and 95% confidence interval for the difference in treatment arms were computed to assess non-inferiority according to the defined margin. A randomization test was used to evaluate the non-inferiority between the treatment arms for the primary safety endpoint (percentage of time below 54 mg dl^{-1}) with a non-inferiority limit of 2%. A two-sided paired t-test was used to evaluate the change in glycated hemoglobin level from baseline to 24 weeks within each arm. A two-sided, two-group t-test was used to evaluate the difference in the mean change of glycated hemoglobin level for the AI-DSS versus physician arms. For both AE count and SAE count, randomization tests were used to test the null hypothesis of no difference between physician and AI-DSS arms. Primary and secondary endpoints were also analyzed in a per protocol cohort that included pairs from AI-DSS and physician arms in which each participant received at least five of the seven scheduled recommendations.

The physician's overrides analysis was computed as the percentage of change from the original AI-DSS recommendations. The recommended daily adjustments were divided into 48 half-hourly periods (half-hour period of available AI-DSS recommendation change) for basal, CR and CF pump setting parameters. There were a total of 144 comparison points (48 time periods and three pump setting parameters) per visit whenever AI-DSS recommendations were available. The recommendation comparison points were compared for the relative changes overall and by the pump setting parameters (that is basal, CR and CF plans).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this Article.

Data availability
Any requests for raw data (that is, glucose levels, insulin delivery, de-identified patient characteristics) will be reviewed by the NextDREAM consortium steering committee, which comprises the principal investigators of the different sites participating in this study. Applications for non-commercial use only will be considered and should be sent to the corresponding author (M.P.). Applications should outline how the specific use of the data would catalyze considerable advancement in the treatment and management of type 1 diabetes or improve care for those living with type 1 diabetes, including the specific purpose of developing therapies and technology that can be used by patients to help manage their disease and improve their health outcomes. Any data that can be shared will need approval from the NextDREAM consortium steering committee and a Material Transfer Agreement in place. All data shared will be de-identified.

Code availability
The code cannot be made available due to proprietary reasons.

References
23. Hahn, S. Understanding noninferiority trials. *Korean J. Pediatr.* **55**, 403–407 (2012).

24. Tamborlane, W. V. et al. Continuous glucose monitoring and intensive treatment of type 1 diabetes. *N. Engl. J. Med.* **359**, 1464–1476 (2008).
25. Aleppo, G. et al. REPLACE-BG: a randomized trial comparing continuous glucose monitoring with and without routine blood glucose monitoring in adults with well-controlled type 1 diabetes. *Diabetes Care* **40**, 538–545 (2017).
26. Abramson, J. H. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. *Epidemiol. Perspect. Innov.* **8**, 1 (2011).

Author contributions

M.P., the corresponding author, had full access to all the data in the trial and takes responsibility for the integrity of the data and the accuracy of the data analysis. R.N., T.B., L.M.L., R.H.S., D.S., S.A.W., K.D., T.D. and M.P. contributed to the study concept and design, contributed data and advised on analysis or interpretation of the data. R.N. and M.P. wrote the first draft of the manuscript with the aid of an independent medical writer. R.N., T.B., L.M.L., R.H.S., D.S., S.A.W., K.D., T.D. and M.P. commented on and revised the manuscript and approved the submission.

Competing interests

R.N. reports receiving grants from Helmsley Charitable Trust, Dexcom and Insulet, personal fees and others from DreaMed Diabetes Ltd, personal fees from Novo Nordisk and Eli Lilly and grants from Medtronic. In addition, R.N. owns DreaMed Diabetes Ltd stock and has a patent licensed by DreaMed Diabetes Ltd. T.B. reports receiving grants and personal fees from Novo Nordisk, Medtronic and Abbott, personal fees from Eli Lilly, Sanofi, Indigo, Roche, AstraZeneca and Dexcom and grants from Glusense and Zealandpharma. In addition, T.B. owns DreaMed Diabetes Ltd stocks. L.M.L. reports receiving grants from Helmsley Charitable Trust and personal fees from Eli Lilly, Novo Nordisk, Sanofi, Convatec, Roche, Insulogic, Dexcom, Insulet, AstraZeneca, Boehringer Ingelheim and Janssen. S.A.W. reports receiving grants from Helmsley Charitable Trust, grants and personal fees from Medtronic and personal fees from Insulet, Tandem, Eli Lilly, Sanofi and Zealand. T.D. reports receiving grants and personal fees from AstraZeneca, DexCom, Boehringer, Novo Nordisk, Medtronic, Sanofi, Eli Lilly and Insulet. In addition, T.D. owns DreaMed Diabetes Ltd stocks. M.P. reports receiving grants from Helmsley Charitable Trust, Dexcom and Insulet, personal fees and others from DreaMed Diabetes Ltd, grants and personal fees from Medtronic and Novo Nordisk, grants from Roche, Eli Lilly and Sanof, grants from Lexicon and OPKO and personal fees from RSP Systems and Qulab Medical. In addition, M.P. owns DreaMed Diabetes Ltd stock and has a patent licensed by DreaMed Diabetes Ltd. R.H.S., D.S. and K.D. declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-020-1045-7.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-020-1045-7.

Correspondence and requests for materials should be addressed to M.P.

Peer review information Jennifer Sargent was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

Outcome*	AI-DSS Arm (n=54)	Physician Arm (n=54)	P value
Mean Glucose (mg/dl)	182 ± 21	180 ± 22	$0.87^{\S}$
SD of glucose (mg/dl)	62 ± 9.9	61 ± 10.7	$0.36^{\S}$
Coefficient of variation in glucose level (%)	35.0 ± 4.0	34.2 ± 4.3	$0.31^{\S}$
Time spent in sensor glucose measurements (%)			
Mean < 50mg/dl	0.9 ± 1.0	0.7 ± 0.7	$0.23^{\dagger}$
Mean < 70mg/dl	3.9 ± 2.7	3.4 ± 2.2	$0.37^{\dagger}$
Mean >180 mg/dl	45.9 ± 11.4	44.9 ± 12.0	$0.67^{\S}$
Mean >240 mg/dl	21.4 ± 9.6	20.4 ± 10.5	$0.76^{\S}$
Mean total insulin dose (U)[¥]	56.9 ± 19.8	52.4 ± 10.7	$0.14^{\dagger}$
Basal insulin dose (U)[¥]	26.5 ± 11	25.9 ± 7.7	$0.74^{\dagger}$
Bolus insulin dose (U)[¥]	29.6 ± 9.7	26.6 ± 8.3	$0.03^{\dagger}$

Extended Data Fig. 1 | Exploratory outcomes: intent to treat cohort (N=108). * Plus-minus values are means $\pm$SD calculated over the entire 24-weeks study period. To convert the values for glucose to millimoles per liter, multiply by 0.05551. This category of SD is an average of the variability of sensor glucose measurements for each patient, rather than the variation in the mean glucose values among patients in the trial. [¥] Insulin information was recorded only for 52 patients per arm. [§] Two-sided ANCOVA adjusting for baseline level and number of valid days. [†] Randomization test.
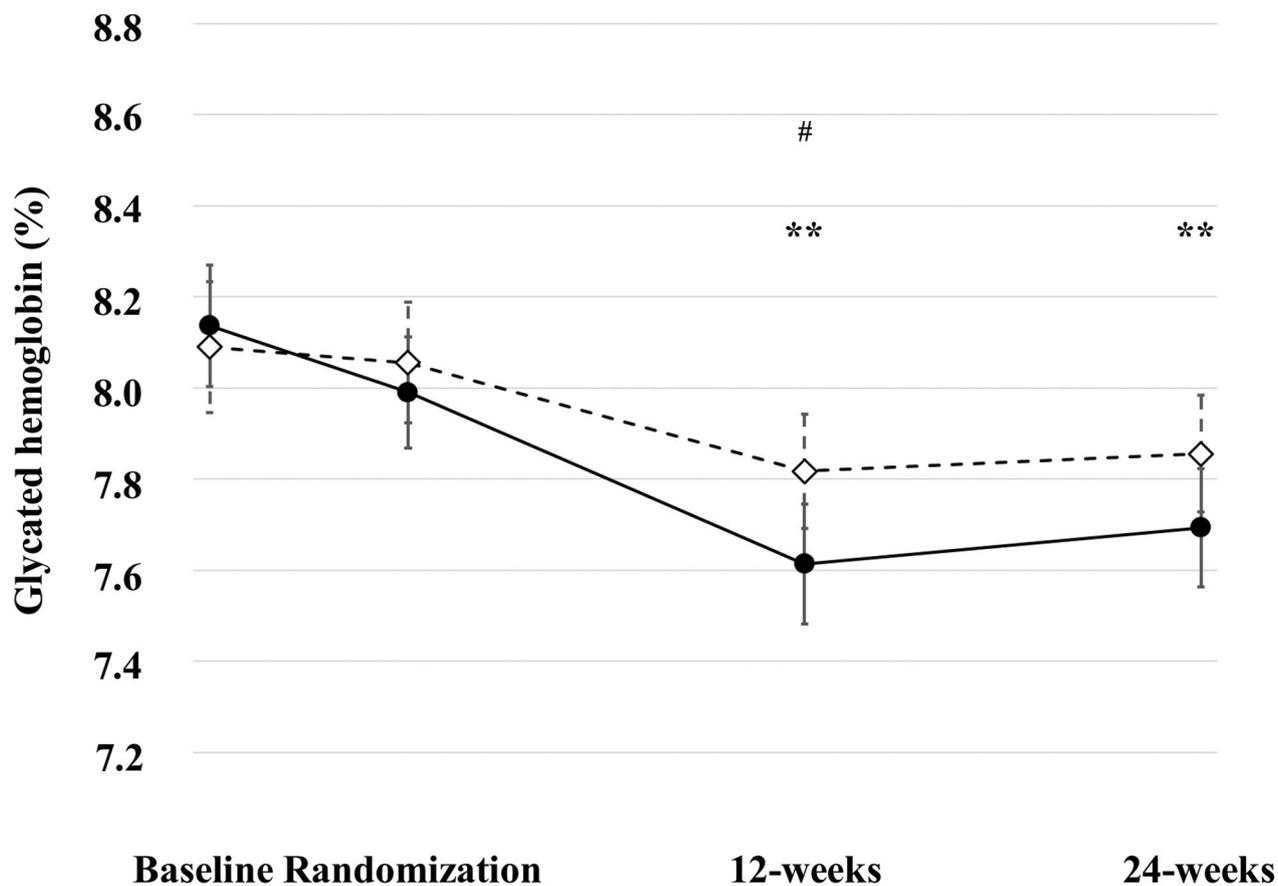
Extended Data Fig. 2 | Glycated hemoglobin levels at baseline, randomization, 12-weeks, and 24-weeks for per-protocol (n=30 per arm) cohort. Values are means (markers) ± standard error (error bars). Black circles mark the AI-DSS arm and empty diamonds mark the physician arm. The results of the two-sided, paired t-test analysis are marked by asterisk symbol to show statistically significant difference between baseline and specific time point within AI-DSS arm and by pound symbol to show statistically significant difference between baseline and specific time point within physician arm, where one symbol denotes P<0.05 and two denotes P<0.01.

Outcome*	AI-DSS Arm (n=30)	Physician Arm (n=30)	P Value
Primary Outcomes			
Time spent with sensor glucose values within target range of 70-180 mg/dl (%)	54.2 ± 10.8	54.6 ± 10.5	$< 1e\text{-}5^{¥}$
Time spent with sensor glucose values below 54 mg/dl (%)	1.0 ± 1.1	1.0 ± 0.8	$< 0.0001^{§}$
Secondary Outcome			
Glycated hemoglobin level			
At Baseline (%)	8.1 ± 0.7	8.1 ± 0.8	
Two-sided 95% CI for the change from baseline to 24-weeks (%)	(-0.74, -0.15)	(-0.57, -0.03)	$0.42^{†}$
Exploratory Outcomes			
Mean glucose level (mg/dL)	175 ± 16.6	175 ± 21.5	$0.83^{Ψ}$
SD of glucose (mg/dL)	58.8 ± 8.8	59.5 ± 10.1	$0.7^{Ψ}$
Coefficient of variation in glucose level (%)	34.0 ± 3.5	34.3 ± 3.7	$0.58^{Ψ}$
Time spent in sensor glucose measurements (%)			
Time below 50 mg/dL	0.7 ± 0.9	0.7 ± 0.6	$0.79^{§}$
Time below 70 mg/dL	3.4 ± 2.3	3.5 ± 2.2	$0.86^{§}$
Time above 180 mg/dL	42.3 ± 10.6	41.8 ± 11.3	$0.83^{Ψ}$
Time above 240 mg/dL	17.7 ± 7.5	17.9 ± 10.1	$0.89^{Ψ}$
Mean total insulin dose (U)	58.2 ± 18.5	54.4 ± 10	$0.18^{§}$
Basal insulin dose (U)	27 ± 12.4	25.2 ± 7.1	$0.47^{§}$
Bolus insulin dose (U)	33.5 ± 12.4	29.4 ± 7.0	$0.14^{§}$

Extended Data Fig. 3 | Study outcomes measurements: per protocol cohort (N=60). * Plus-minus values are means ±SD. To convert the values for glucose to millimoles per liter, multiply by 0.05551. ¥ One-sided test using ANCOVA model adjusted for baseline % of time in range and number of valid days used in the analysis. § Non-inferiority randomization test. †Two-sided paired t-test. ΨTwo-sided ANCOVA adjusting for baseline level and number of valid days.

Items	Week 12 (n = 8)	Week 24 (n = 13)
Using the Advisor Pro was intuitive and simple	4.5	4.8
I found the Advisor Pro to be reliable	4.1	4.5
I believe the Advisor Pro was safe	4.5	4.4
The Advisor Pro was accurate (i.e. the algorithm worked as it should when providing advice)	4.3	4.3
I trusted the Advisor Pro and the information it provided	4.1	4.3
The Advisor Pro saved me time	4.3	4.3
I always took the Advisor Pro advice	4.1	4.2
The Advisor Pro was similar to therapy adjustments I would have done clinically	3.3	3.5
The Advisor Pro suggested better adjustments than I would have done clinically	2.9	3.1
Overall, my experience of using the Advisor Pro met my expectations	4.1	4.5
Overall, my experience of using the Advisor Pro was positive	4.3	4.5
I believe using the Advisor Pro positively impacted on the self-management actions of my patients	4.0	4.4
The Advisor Pro was useful in helping me communicate insulin dosing decisions to my patients	4.6	4.5
The Advisor Pro was sufficiently dynamic to provide accurate advice in different situations	4.3	4.2
The Advisor Pro helped remove some of the uncertainty from decision-making and improved my overall clinic experience	3.3	3.6
The training materials for the Advisor Pro were easy to understand	4.1	4.5
The process of sending patient data to the Advisor Pro was easy	3.7	4.2
The process of receiving data from the Advisor Pro was easy	3.6	4.1
The Advisor Pro saved me time in analyzing a patient's data	4.1	4.2
The suggested Advisor Pro insulin adjustments were clear	4.3	4.5
The Advisor Pro reasoning for the suggested therapy adjustments was clear	3.8	4.2
The Advisor Pro reasoning for the suggested therapy adjustments was sound	3.8	4.2
The Advisor Pro reinforced my own clinical assessment	3.6	3.8
The Advisor Pro allowed for more frequent changes, getting to optimal dosing more quickly	4.1	4.0
The Advisor Pro helped me identify patterns difficult to identify	4.0	3.8
The Advisor Pro allowed me to make less aggressive insulin adjustments	3.0	3.2
The Advisor Pro allowed me to make more aggressive or larger insulin adjustments	3.5	3.4
The Advisor Pro allowed me to recognize my patient's behavior in self-managing diabetes	3.8	3.7

Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Healthcare professionals post-intervention survey responses: physicians who used the AI-DSS during the study (n=13).
The scale of the reply was: 5—Strongly Agree, 4—Somewhat Agree, 3—Do not Agree or Disagree, 2—Somewhat Disagree, 1—Strongly Disagree. The Healthcare Professionals Post-Intervention Survey is a 50 items questionnaire. The questionnaire comprises: (a) 28 items for pertaining to the physician's experience with the Advisor Pro use and recommendations. Each item is score on a 5-point scale range from 5 (strongly agree) to 1 (strongly disagree) (Extended Data Fig. 4) (b) 22 items are questions that asses the physician's view regarding integration of the Advisor Pro into daily routine practice (14 items are yes/no questions and 8 items are open questions). The questionnaire was developed by Prof Katharine Barnard PhD CP sychol AFBPsS and reviewed by the investigators (first and third authors).

	AI-DSS Arm (n=54)	Physician Arm (n=54)
# of documented insulin titration visits (i.e. from total of 7 insulin titration visits per patient, equal to 378 insulin titration visits per arm).	367	374
# of given recommendations (% from # of insulin titration visits)	275 (74.9%)	328 (87.7%)
# of recommendation by the AI-DSS (% from given recommendation)	268 (97.4%)	

The means in which recommendations were shared with the participents[1]

	AI-DSS Arm (n=54)	Physician Arm (n=54)
# of recommendations shared via the DMS (% from given recommendations)	249 (90.5%)	282 (86.0%)
# of recommendations shared via Emails (% from given recommendations)	16 (5.8%)	7 (2.1%)
# of recommendations shared via Phone (% from given recommendations)	6 (2.2%)	19 (5.8%)
# of recommendations provided in person/ visit (% from given recommendations)	2 (0.7%)	20 (6.1%)
Unknown (% from given recommendations)	2 (0.7%)	0 (0%)

Reason for missing scheduled recommendation[2,3]

	AI-DSS Arm (n=54)	Physician Arm (n=54)
Technical issues[4] (% from total missing recommendations)	22 (22.2%)	10 (21.7%)
Insufficient data[5] (% from total missing recommendations)	67 (67.7%)	31 (67.4%)
Pump clock shifting[6] (% from total missing recommendations)	5 (5.1%)	3 (6.5%)
Unknown (% from total missing recommendations)	5 (5.1%)	2 (4.3%)

Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Description of the number and technical issues related to insulin dose recommendations in both study arms - intent to treat cohort. 1 In both groups, there were cases in which insulin dosing recommendation was not shared with the participant via the DMS system but in another method. 2 The reasons for missing scheduled recommendations were: physician did not give a recommendation for the physician arm (supplement to the number of given recommendation) and lack of minimal requirements needed for creating advisor recommendation such as lack of valid days for analysis for the AI-DSS arm (supplement to the number of recommendation by the AI-DSS). 3 In 7 cases, recommendation was given by the physician and not the Advisor. In 5 of these cases the minimal data requirement of the AI-DSS was not meet and in the other 2 cases there was a suspected pump clock shifting that prevented the system from providing a recommendation. 4 Technical issues that caused missing recommendation include: any technical failure related to pump, continuous glucose monitoring, diabetes management system or temporary algorithm faults that were resolved during the study. 5 Insufficient insulin and glucose data according to the system requirements prevent formation of recommendation in both trial arms. 6 Pump clock shifting is a safety layer of the advisor that prevent the formation of a recommendation.
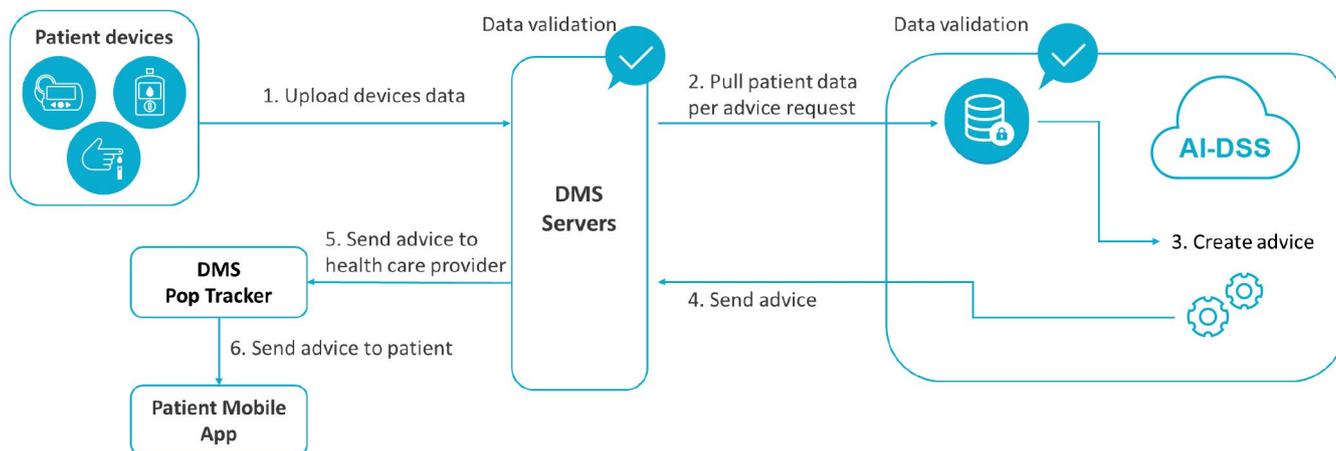
Extended Data Fig. 6 | AI-DSS data flow in the study. 1. participants upload devices data to DMS 2. AI-DSS pulled data from DMS after DMS requested advice and validated data 3. AI-DSS detected the source of the data: in case the data came from participants in the physician arm no advice was created and in case the data came from participants in the AI-DSS arm advice was created 4. AI-DSS report was sent via DMS to the physician 5. Physician could review new pump settings recommendations within AI-DSS report page using the DMS platform 6. Physician shared new pump settings and management tips with participant through smartphone.
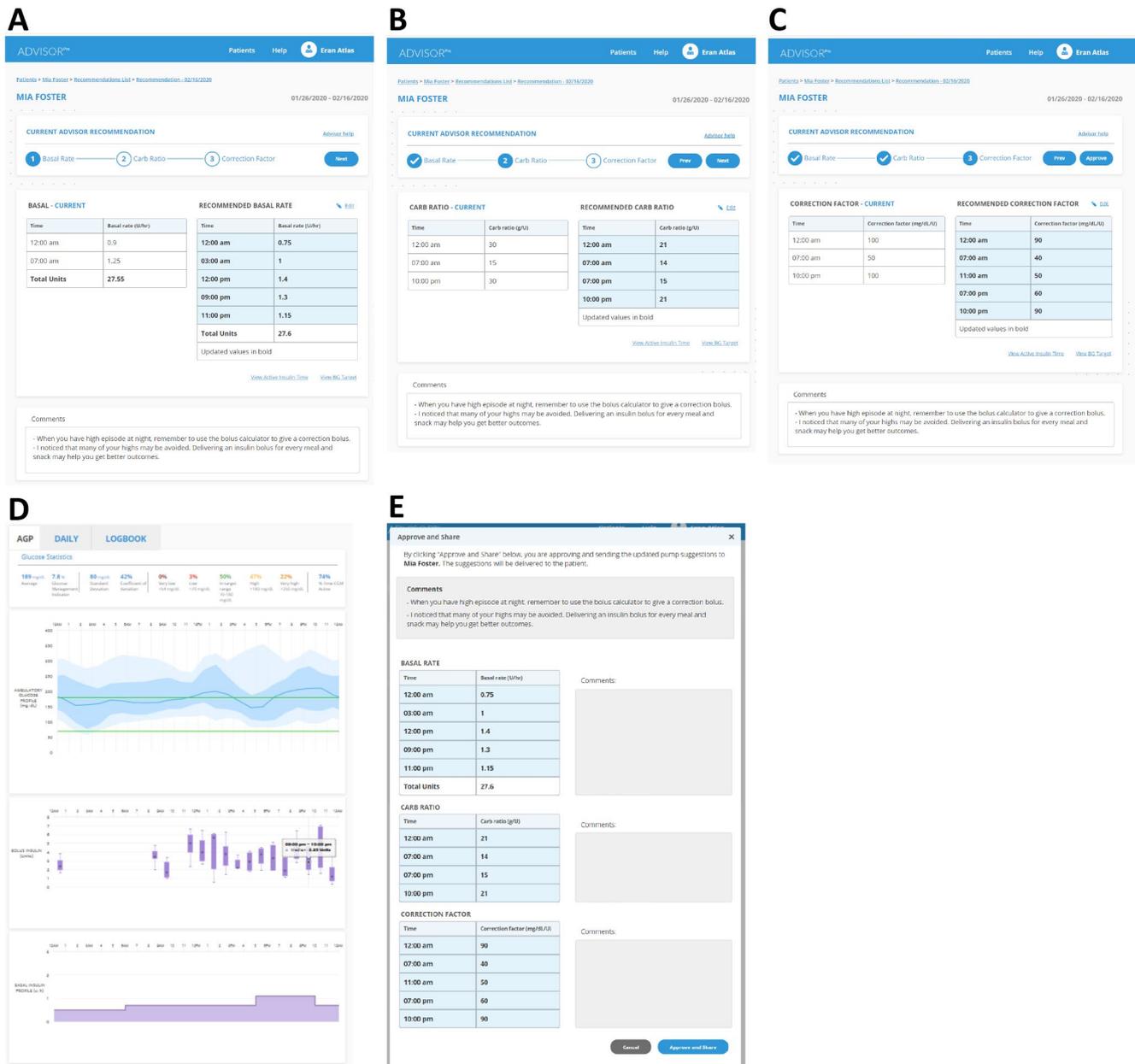
Extended Data Fig. 7 | Data management system population tracker with AI-DSS report. a, Once the participant uploaded data to the DMS and algorithm has enough data, a new recommendation will be sent and physician could enter into the patient page and review it. Current basal rates are displayed on the left, for the physician arm—a duplicate of the current settings was displayed on the right, clinicians could click EDIT and make changes if necessary and add their own comments for behavioral tips. For the AI-DSS arm—the right side displayed the AI-DSS recommended changes (if necessary) in bold. Comments may be added from the AI-DSS. **b,** Current carbohydrate ratio values are displayed on the left, for the physician arm - a duplicate of the current settings was displayed on the right, clinicians could click EDIT and make changes if necessary and add their own comments for behavioral tips. For the AI-DSS arm—the right side displayed the AI-DSS recommended changes (if necessary), in bold. Comments may be added from the AI-DSS. **c,** Current correction factor (Insulin Sensitivity Factor) are displayed on the right, for the physician arm - a duplicate of the current settings was displayed on the right, clinicians could click EDIT and make changes if necessary and add their own comments for behavioral tips. For the AI-DSS arm—the right side displayed the AI-DSS recommended changes (if necessary), in bold. Comments may be added from the AI-DSS. **d,** The Ambulatory Glucose Profile, Logbook and Daily reports are displayed under each of the previous settings screens for guidance. **e,** Once all settings screens are completed, a summary of the recommendation will appear. This is what the patients sees on their mobile app once the physician shares the recommendations with them.

INSULIN PUMP SETTINGS RECOMMENDATION

Report from 05/31/2020 - 06/21/2020

On 07/23/2020 your provider recommended that you change your pump settings to the following values:

General Comments

 - You are over treating your lows. Eat moderately when treating your lows.

NEW BASAL PLAN SETTINGS

Time	Basal rate (Unit/hr)
12:00 AM	2
10:00 AM	1.5
9:00 PM	2
Total Units	**42.5**

NEW CARB RATIO PLAN SETTINGS (IC RATIO)

Time	Carb ratio (g/Unit)
12:00 AM	10
6:00 AM	12
6:00 PM	6

No comments

NEW CORRECTION FACTOR SETTINGS PLAN (INSULIN SENSITIVITY)

Time	Correction factor (mg/dL/Unit)
12:00 AM	50
10:00 PM	120

No comments

Extended Data Fig. 8 | AI-DSS insulin pump recommendations presented on patient's smartphone: new basal plan, new carbohydrates ratio, new correction factors and diabetes management tips. The recommendation report for the patient is presented in the app, three screen shots from the application. The data in the figure related to virtual patient, Mia Foster.

Insulin Dose Adjustments
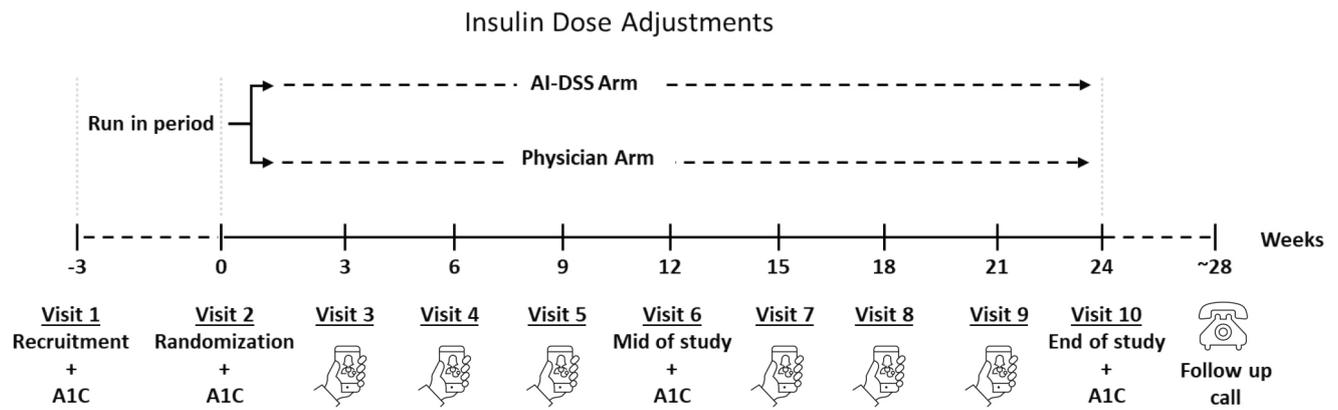
Extended Data Fig. 9 | Study design scheme. After a 3-week run-in period, participants were randomized to participate either at the IA-DSS or physician arm. Three weeks after randomization and every 3 weeks thereafter, all participants uploaded their insulin pump and continuous glucose monitoring data using the DMS. Participants then received insulin pump settings recommendations either from physician or AI-DSS according to their randomization for a 24-week period.

Screening Week -3	• Written informed consent obtained and eligibility assessed. • Medical history (concomitant medication, concomitant illness, known allergies, demographic information, and current insulin treatment (insulin type, dosages) were documented and anthropometric characteristics and vital signs were assessed. • Capillary blood samples were obtained for on-site measurement of baseline Hb1c and urine pregnancy testing was performed for all female patients of child-bearing age. • Eligible patients underwent an educational session on CGM use and general diabetes guidelines, including sensor calibration, technical issues, problem-solving and safety issues. Participant competency in using CGM was evaluated by the study nurse.
Run-in Period Weeks -3 to 0	• Participants who regularly used continuous glucose monitoring, were asked to continue with their treatment. Participants who used only SMBG for glucose monitoring were asked to use CGM for the remainder of the study. • All patients were asked to use fingerstick testing to measure their capillary glucose at least 4 times per day, prior to meals and at bedtime, and continue to use their insulin pump for insulin delivery.
Randomization Week 0	• Insulin pump, CGM and SMBG data were uploaded and analyzed and adverse events were recorded. • Participants completed the Diabetes Treatment Satisfaction Questionnaire status version (DTSQs). 14 • Participants were randomly assigned in a 1:1 ratio to either the intervention group (advisor) or control group (expert physician arm). • Participants and study clinicians underwent a training session on how to download the insulin pump, CGM and blood glucose meter data at home. • Participants were instructed to download their data via the DMS platform every 3 weeks for the first 3 months.
Phone Visits Weeks 3, 6, 9, 15, 19, 21	• Phone visits were conducted to confirm that the recommended pump adjustments sent to patients had been correctly entered into their insulin pumps. • Adverse events and any changes in concomitant medications were recorded. • All advisor recommendations, recommendations overridden by the physician/prescriber and reason for overriding the recommendation were documented.
Clinic Visits Weeks 12, 24	• Data from patients' insulin pump, CGM and glucose meter were uploaded. • Adverse events and changes in concomitant medications were recorded. • HbA1c was assessed and a urine pregnancy test was performed with female patients with childbearing potential. • Participants completed the DTSQs. • Clinicians completed the Healthcare Professional Post-Intervention Survey. (Weeks 12 and 24)

Extended Data Fig. 10 | Study visit schedule. Procedures conducted at each visit.

nature research

Corresponding author(s):   Moshe Phillip

Last updated by author(s): 2020/07/23

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	A CRO company (Clinlogix) was responsible for data collection and management from all 7 participating sites. Glucose and insulin data was available from the participants continuous glucose monitoring uploads (all participants used the Dexcom G5 Device). Uploads were done to Glooko diabetes management platform software version 19.1-19.2 for web site, and Mobile Glooko version 5.3-5.4.
Data analysis	Data were analyzed using the R statistical environment, version 3.6.1. Dedicated code was written to carry out the analyses using routines that are available in the software and a specialized routine for performing statistical tests via randomization. The sample size was calculated using WINPEPI, computer program for epidemiologists version 11.39.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Any requests for raw data (i.e. glucose levels, insulin delivery, de-identified patient characteristics) will be reviewed by the NextDREAM consortium steering committee which comprises of the principle investigators of the different sites participated in this study. Applications for non-commercial use only will be considered and should be sent to the corresponding author (MP). Applications should outline how the specific use of the data would catalyze considerable advancement in the treatment and management of type 1 diabetes or improve care for those living with type 1 diabetes; as well as, including the specific purpose

of developing therapies and technology that can be used by patients to help manage their disease and improve their health outcomes. Any data that can be shared will need approval from the NextDREAM consortium steering committee and a Material Transfer Agreement in place. All data shared will be de-identified.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the sample size calculation, data from the JDRF CGM randomized controlled study (ref.15) and the REPLACE-BG study (Ref. 16) were used to estimate properties of the primary endpoint and to define the non-inferiority limit. The sample size was calculated using WINPEPI17 for non-inferiority analysis with a margin of 7.5% between arms in a parallel study design with power of 90%, a one-sided significance level of 5% and SD of 13% with a correlation of 0.48 between the baseline and outcome time in range over the 24-week treatment period. The resulting analysis showed that 104 subjects were needed for the study. The sample size was set at 112 participants (16 patients at each site), given an estimated 5-10% drop-out rate, with randomization in a 1:1 ratio between the two study arms. A detailed description available in the article statistical analysis section
Data exclusions	5 participants were excluded due to randomization that was not according to the designed site A1c strata for randomization. In addition 9 participants had lack of data or valid pair for analysis. See Figure 2 (Study Flow Chart). The randomization strata and the definition for valid pair were pre-established. Participants were paired within strata within site, the analyses that compare the AI-DSS and physician groups was predefined to be limited to the 56 pairs specified in the original design and also to those pairs for which follow-up data were available for both members of the pair.
Replication	The study design and data analysis according to the provided protocol can be repeated.
Randomization	Eligible participants were randomly assigned in a 1:1 ratio to either the Advisor arm or Physician arm, using computer-generated permuted block randomization stratified by site, age (10-14 years, 15-17 years, 18-21 years) and glycated hemoglobin level at screening (7-<8% [53-<64 mmol per mol], 8-<9% [64-<75 mmol per mol], and 9-10% [75-86 mmol per mol]) to achieve equal distribution of participants in the two study arms. Each site was assigned 8 of the 9 age-glycated hemoglobin strata and asked to recruit one pair of participants in each stratum (e.g., each pair had one participant designated to AI-DSS arm and the other to the Physician arm).
Blinding	The investigators were not blinded to the group allocation as each investigator had to give pump settings adjustments to each participant according to the participant designated allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	*Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Validation	*Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Human research participants

Population characteristics	Children, adolescents, and young adults with type 1 diabetes for at least one year and treated with insulin pump therapy were eligible for participation in the study. Average age was 15.6 ± 3 years and 52% were females.
Recruitment	Participants were recruited at each participating academic center - 7 clinical sites; 4 in the USA (University of Florida; Joslin Diabetes Center; Barbara Davis Center; Yale University School of Medicine), 2 in Europe (University Children's Hospital Ljubljana Slovenia, Children's Hospital AUF DER BULT, Germany), and one in Israel (Schneider children's medical center of Israel).Eligible patients were offered to participate in the study while arriving to routine clinical visit and those who agreed were enrolled to the study. Selection bias may occurred due to some of the eligibility criteria, mainly: Bone Mass Index-SDS – below the 97th percentile for age, participants/ parents were required to have minimum computer skills and understanding of navigating the Internet as well as to have a smartphone (Apple or Android and Windows). Therefore, we evaluate that the study results cannot be generalized to patients above the 97th percentile for age and who do not have smartphone and minimal computer skills.
Ethics oversight	The study protocol was approved by the local institutional review boards at each participating center as follows: Joslin Diabetes Center, Committee of Human Studies; Barbara Davis Center for Childhood Diabetes, Colorado Multiple Institutional Review Board;Yale University School of Medicine, Yale University Human Investigation Committee; Diabetes Institute University of Florida, Institutional Review Board University of Florida ; Schneider children's medical center of Israel, Institutional Review Board Rabin Medical Center; University Children's Hospital Ljubljana Slovenia, national regulatory authorities- Republic of Slovenia National Medical Ethics Committee (JAZMP); Children's Hospital AUF DER BULT, national regulatory authority - Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). All participants (or parents/legal guardians) provided written informed consent, and assent according to local requirements.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

All manuscripts should comply with the ICMJE <u>guidelines for publication of clinical research</u> and a completed <u>CONSORT checklist</u> must be included with all submissions.

Clinical trial registration	ClinicalTrials.gov number, NCT03003806
Study protocol	Study protocol was uploaded with the manuscript submission - available as supplementary material
Data collection	A CRO company (Clinlogix) was responsible for data collection and management from all 7 participating sites. Clinlogix provided the data to the external statistician. Glucose and insulin data from the continuous glucose monitoring device was generated from Glooko system uploads by DreaMed Diabetes for analysis by an external statistician. Participants were recruited between 20 November, 2017 first patient first visit (FPFV) and 10 December, 2018 last patient first visit (LPFV) over 12 months period. Data collection was performed throughout the entire study period between November 2017 enrollment of first the participant (FPFV), until August 2019, two weeks after Last Patient last Visit was performed. Data were collected at each participating site using electronic case report form.
Outcomes	The primary efficacy outcome was the percentage of time that the glucose level, as measured by the continuous glucose monitor, was in the target range of 70 to 180 mg per deciliter (3.9 to 10.0 mmol per liter) over the active treatment period. The primary safety outcome was the percentage of time that the glucose level was less than 54 mg per deciliter (3.0 mmol per liter). Secondary outcomes were changes in glycated hemoglobin level from baseline to the end of study and adverse events and severe adverse events. An adverse event was defined as any untoward medical occurrence in a subject and does not imply that there is a relationship between the adverse event and the device under investigation. An adverse device event was defined as adverse event related to the use of an investigational medical device. Severe adverse event was defined as any adverse event occurring in a clinical trial or in a performance evaluation needing approval that, directly or indirectly, has led, could have led or could lead to death or a serious deterioration in health condition of a subject, a user or any other person, without considering whether the event was caused by the medical device. Severe device adverse event defined as effect that has resulted in any of the consequences characteristic of a serious adverse event. The primary safety and efficacy outcomes were assessed from the continuous glucose monitoring device uploads (all participants used the Decome G5 continuous glucose monitoring) . The secondary outcome was glycated hemoglobin that was assessed at each participating site using capillary A1c measurement. Adverse events were collected during the study from each study site at the electronic case report form.