SUPPLEMENTAL METHODS

Study design and participants

Exclusion criteria were: planning a trip likely to cause jet lag; planning a hospital admission within 3 months; suffering from serious illness including end-stage kidney disease, uncontrolled psychiatric disorders, and arrhythmia; taking systemic glucocorticoids or immunosuppressive drugs; and potential pregnancy. Individuals inexperienced in using wearable devices or smartphones were also excluded. Among the 26 subjects screened for participation, two were excluded: one was diagnosed with type 1 diabetes and the other had a poor ability to handle wearable devices.

Protocol

Patients were informed that poor compliance could result in withdrawal from the study. Poor compliance was defined as going more than 5 consecutive days without wearing the device during the screening period of 2 weeks. Trained nurses measured anthropometric data, including height, body weight, and waist circumference. Individual medical history, demographic information, and prescribed drugs were recorded at baseline and follow-up visits. Laboratory variables, including hemoglobin A1c (HbA1c), fasting plasma glucose, serum creatinine, and lipid profiles, were also evaluated. Anthropometric measurements and laboratory tests were repeatedly performed at the end of the study period. During the study period, patients' glucose-lowering agents were left unchanged.

Wearable device-derived measures

First, step count indicated the mean step count for a day during the study period. The higher the mean step count, the higher the physical activity. Second, the heart rate (HR) was measured every 5 seconds using an activity tracker. The resting HR indicates the average HR within timeslots when a user has no activity. Consecutive HR data at 48-hour intervals were analyzed with cosine curve fitting to reflect circadian rhythm, which provided four representative parameters: (1) HR_consistency, (2) HR_magnitude, (3) HR_misalignment, and (4) HR_variability. The higher the mean of the first two parameters (HR_consistency and HR_magnitude), the more regular the HR of the patients. The higher the mean of the last two parameters (HR_misalignment and HR_variability), the more irregular may be the circadian rhythm of the patients. A higher standard deviation (SD) of each parameter indicates a greater fluctuation of the circadian rhythm. Lastly, for sleep-related parameters, we regarded 'bedtime' as a constant time of sleep as the period from eight hours (ideal sleeping time) before sunrise until sunrise the next day. To evaluate the regularity of the sleep cycle, variables regarding the time differences in going to sleep and waking up were created. Other data on sleep patterns such as sleep duration and quality were also evaluated. A more detailed description of each of the variables mentioned above is provided in Supplemental Table S1.

Statistical analysis

In our analysis, we used the complete dataset of 1,247 days (mean±SD days per patient, 52.0±29.1) from 24 patients through a process of creating step-, HR-, and sleep-related features. To describe the baseline characteristics, data were presented as a number (percent) for categorical variables and a mean±SD for continuous variables. The Shapiro-Wilk test was used to assess the normality of the distribution of variables. To compare continuous variables between Groups A and B, the independent *t* test and Wilcoxon rank-sum test were used, as appropriate.

Regarding the Shapley additive explanation (SHAP) values, these represent a powerful system of evaluating feature importance in predictive models with consistent results [8,17]. SHAP dependence plots were evaluated for the top 20 important variables to examine how the importance of these features changed with respect to themselves and how much each feature contributed, either positively or negatively, to coefficient values in logistic regression [8,17].

SUPPLEMENTAL RESULTS

Subjects in Group A were likely to be older (50.1±6.8 vs. 44.9±7.7, P=0.106), have a longer duration of diabetes (11.9±4.5 years vs. 5.7±3.7 years, P=0.001) and a lower body mass index (25.4±4.4 kg/m^2 vs. 29.5±5.4 kg/m^2, P=0.071) compared to those in Group B. The mean HbA1c level was similar between groups (7.4% vs. 7.3%), but the mean total cholesterol level (120.1±20.6 mg/dL vs. 159.4±35.1 mg/dL, P=0.006) and low-density lipoprotein cholesterol level (66.2±18.3 vs. 88.2±29.2, P=0.056) were lower in Group A than in Group B. Subjects in Group A also had healthier habits such as a lower rate of current smokers, alcohol consumption, and a lower number of shift workers.