Comparing Real-Time Self-Tracking and Device-Recorded Exercise Data in Subjects with Type 1 Diabetes

Danielle Groat1,2 Hyo Jung Kwon2 Maria Adela Grando2,3 Curtiss B. Cook2,3 Bithika Thompson3

1 Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah, United States
2 Department of Biomedical Informatics, Arizona State University, Scottsdale, Arizona, United States
3 Division of Endocrinology, Mayo Clinic Arizona, Scottsdale, Arizona, United States

Address for correspondence Danielle Groat, PhD, 421 Wakara Way Suite 140, Salt Lake City, UT 84108, United States (e-mail: danielle.groat@utah.edu).

Abstract

Background Insulin therapy, medical nutrition therapy, and physical activity are required for the treatment of type 1 diabetes (T1D). There is a lack of studies in real-life environments that characterize patient-reported data from logs, activity trackers, and medical devices (e.g., glucose sensors) in the context of exercise.

Objective The objective of this study was to compare data from continuous glucose monitor (CGM), wristband heart rate monitor (WHRM), and self-tracking with a smartphone application (app), iDECIDE, with regards to exercise behaviors and rate of change in glucose levels.

Methods Participants with T1D on insulin pump therapy tracked exercise for 1 month with the smartphone app while WHRM and CGM recorded data in real time. Exercise behaviors tracked with the app were compared against WHRM. The rate of change in glucose levels, as recorded by CGM, resulting from exercise was compared between exercise events documented with the app and recorded by the WHRM.

Results Twelve participants generated 277 exercise events. Tracking with the app aligned well with WHRM with respect to frequency, 3.0 (2.1) and 2.5 (1.8) days per week, respectively (p = 0.60). Duration had very high agreement, the mean duration from the app was 65.6 (55.2) and 64.8 (54.9) minutes from WHRM (p = 0.45). Intensity had a low concordance between the data sources (Cohen’s kappa = 0.2). The mean rate of change of glucose during exercise was −0.27 mg/(dL min) and was not significantly different between data sources or intensity (p = 0.21).

Conclusion We collated and analyzed data from three heterogeneous sources from free-living participants. Patients’ perceived intensity of exercise can serve as a surrogate for exercise tracked by a WHRM when considering the glycemic impact of exercise on self-care regimens.

Keywords
► exercise
► smartphone app
► wearables
► type 1 diabetes
► patient-generated data

Background and Significance

Type 1 diabetes (T1D) is a complex, chronic disease which requires patients to engage in certain self-management strategies to avoid acute and long-term complications.1–3 For instance, adherence to insulin therapy, medical nutrition therapy, and physical activity are the ways and means by which individuals with T1D achieve the objective of optimizing glycemic control.4–6 Patients are more likely to adhere to self-management regimens that incorporate personal lifestyle...
preferences. However, incorporating physical activity can be challenging for individuals with T1D due to the multiple factors that influence glucose levels, such as the duration and intensity of exercise, the timing and quantity of delivered insulin, and carbohydrates consumed. Recommendations for compensating for exercise include disconnecting from the pump, adjusting bolus and/or basal insulin, consuming carbohydrates, and waiting for glucose levels to be at an appropriate concentration. It is also suggested that the current blood glucose concentration and the duration and intensity of the planned exercise should be considered before engaging in exercise to employ the appropriate compensation techniques to offset the acute effects of physical activity. Studies have found that most individuals with T1D resort to trial and error as they compensate for exercise, which can lead to considerable variability in self-care behaviors between and within participants.

Ideally, clinicians would be able to review patient-generated exercise logs and compensation techniques employed, along with glucose data, to provide personalized recommendations aimed at improving glycemic control. However, one of the challenges in individualizing recommendations for exercise is the difficult nature in establishing the individual impact of physical activity on glucose levels in real-world, noncontrolled environments. The rate of change of glucose levels across time (ΔG/min), is an indicator of how aggressively one should compensate for exercise, and has been found to be related to the intensity of physical activity, with moderate exercise decreasing glucose levels at a faster rate when compared with light exercise, and vigorous exercise sometimes increasing or decreasing glucose levels. Commonly used techniques for measuring physical activity include device-recorded data (e.g., pedometers or heart rate monitors), or self-reported information. The accuracy of self-reported exercise data has mixed levels of validity, while devices may fail to produce accurate measurements.

Wearable activity monitors have been found to have good levels of reliability with respect to step counts and distance, with other measurements such as energy expenditure and sleep quality having lower validity. While one research group used a wrist-worn activity monitor to validate self-reported exercise in heart failure patients prescribed a walking routine, due to the shortcomings of these methods, there is no established standard for measuring exercise outside of controlled research settings. Although it has been suggested that combining methods to measure exercise in real-life outpatient settings may be the best way to accurately assess periods of exercise, this approach remains understudied.

Studies have found that one-quarter to one-third of patients with T1D use apps to provide assistance with carb counting, recording glucose levels, and tracking exercise, with insulin pump users more likely to use apps than those on multiple daily injections. The authors have previously published data on a noncommercial smartphone application (app) called iDECIDE. To the best of our knowledge, iDECIDE is the first app that allows users to track the techniques they employ to compensate for meals, alcohol, and exercise, as illustrated in Fig. 1. Additional functionalities of the app include self-tracking exercise duration, intensity, and start time. Outside of the scope of this study are features for self-tracking carbohydrates for meals and alcohol, assisting with carbohydrate counting, and providing insulin bolus recommendations. Utilization of the app in previous studies on T1D patients on insulin pumps has identified unexpected self-management behaviors related to how patients compensate for variables such as exercise. For instance, subjects reported on a survey that they adjusted the basal rate when exercising, but they recorded with the app that they instead took snacks before exercising.

The objective of this study was to compile and compare exercise behaviors from free-living participants with T1D from three heterogeneous data sources found in outpatient settings to answer two questions: (1) how does an individual’s exercise behaviors self-tracked via the app compare with real-time recording by a wristband heart rate monitor (WHRM)?, and (2) how does ΔG/min attributable to exercise tracked with the app, compare with ΔG/min attributable to exercise recorded by a WHRM?

Methods

Participant Recruitment

After obtaining Institutional Review Board approval, we recruited patients from an outpatient academic endocrinology clinic who were 18 to 70 years old and who were being treated for T1D. All participants used a Medtronic (Minneapolis, Minnesota, United States) insulin pump paired with CGM, and all owned a smartphone. Patients in fragile health, limited life expectancy, a history of mental health problems, advanced vascular disease or microvascular complications, and known history of severe hypoglycemia were excluded. Participants were required to have been receiving care in the clinic for at least 1 year. Patients were identified by facility researchers during routine office visits after which a recruitment appointment was set where the participants gave informed consent.

Exercise Data

At the recruitment appointment, participants were instructed to maintain their normal routines for 1 month. They were asked to log their exercise activities by using the self-tracking module of the app, which was installed on each participant’s smartphone. Participants tracked exercise by recording perceived intensity as indicated by the “talk test” as: light (able to talk and sing), moderate (able to talk but not
The talk test has been found to be a valid and reliable tool for monitoring exercise intensity across individuals of various health status. Additionally, exercise duration and start time were tracked (Fig. 1). At the end of the month, logs were downloaded in tabular format from the cloud-based Web service.

Simultaneously with the use of the app, an off-the-shelf WHRM which did not require any additional steps for calibration (FitBit Charge HR, San Francisco, California, United States) was distributed to each participant and placement was encouraged on the nondominant hand. The WHRM recorded heart rate in beats per minute, and start time and duration of exercise, accurate to the minute. During exercise events, a heart rate reading was displayed every 30 seconds, and an overall average heart rate for the event was also provided by the device manufacturer’s Web portal, otherwise a heart rate reading was provided every 5 minutes. The WHRM recorded exercise in real-time fashion for the 1-month study period and was synchronized to the manufacturer’s companion app for data uploads to the Web portal which granted researchers access to the data.

Glucose Readings
Insulin pumps are capable of synchronizing and storing data from blood glucose meters and CGM. CGM delivers an interstitial glucose measurement to the insulin pump every 5 minutes. Participants were instructed to engage in self-care as usual during the study, and no additional instructions were given regarding the calibration of the CGM sensor. Participants uploaded their insulin pump and CGM data to the manufacturer’s Web portal and 1 month of corresponding data were downloaded in raw tabular format. Readings from the CGM were used to calculate $\Delta G/min$.

Data Analysis
Comparison of Exercise Behaviors between the App and WHRM
Exercise events self-tracked with the app and recorded by the WHRM were categorized as occurring in both data sources (i.e., matched), or as being present in only one data source (i.e., unmatched). Exercise events were considered to occur in both data sets if portions of the exercise event from both data sources occurred within a 1-hour window. Exercise parameters (frequency, duration, intensity) from the app and the WHRM were compared. The average number of days per week participants engaged in exercise was determined for both the app and the WHRM. Exercise duration was also...
compared for exercise events that matched in both sources of data.

The intensity of exercise tracked by the participants with the app (light, moderate, or vigorous) was compared against the WHRM data. Although the heart rates from the WHRM were categorized into three intensities according to the manufacturer’s proprietary algorithms, we categorized the average and highest heart rate recorded by the WHRM based on the formula: \( \text{intensity} = \frac{\text{heart rate}}{(208 - 0.7 \times \text{age})} \). The intensity was categorized as vigorous exercise for values between 0.7 and 0.85, moderate from 0.5 to 0.7, and light from 0.3 to 0.5. The calculations were performed with age tracked in years, and heart rate in beats per minute.

**Comparison of \( \Delta G/\text{min} \) between the App and WHRM**

A start and finish glucose reading were automatically extracted from the insulin pump data for exercise events from the app and the WHRM. Exercise events were included only if both a start and finish glucose value from the CGM within 10 minutes of the documented start and finish time for exercise were available. The \( \Delta G/\text{min} \) was calculated as \((\text{finish glucose} - \text{start glucose}) / \text{exercise duration}\), where finish glucose and start glucose values were reported in mg/dL and exercise duration was in minutes. The \( \Delta G/\text{min} \) for exercise tracked with the app was compared with the \( \Delta G/\text{min} \) for exercise documented by the WHRM.

**Statistical Analysis**

Cohen’s kappa was used for categorical comparisons and two-sided t-tests of unequal variance were used for numerical data, with pairing used for matched exercise events. Single-factor analysis of variance was used to compare \( \Delta G/\text{min} \) categorized by intensity between the app and the WHRM. All results are reported as mean (standard deviation).

**Results**

**Participant Characteristics**

Twelve patients with T1D were recruited, all were Caucasian. All of the eight females and one male engaged in exercise and contributed complete data. One male participant did not engage in physical activity and was removed from any further analysis. Another male participant did not track any exercise with the app, but the WHRM recorded exercise events, while another male participant tracked exercise with the app, but was unable to upload exercise data from the WHRM to the manufacturer’s companion portal. The average age of the 11 participants that generated exercise data was 48 (13.4) years and average length of T1D diagnosis was 29 (12.5) years. One participant used the Paradigm Revel 723 insulin pump paired with a Dexcom CGM, and all others used the MiniMed 530G pump paired with the Enlite CGM. Participants had an average of 13 (6.1) years of insulin pump therapy and average hemoglobin A1c (HbA1c) of 7.7% (1.0%).

There were 161 exercise events tracked with the app from 10 participants. The WHRM recorded 116 exercise events from the 10 participants that uploaded data. There were 277 exercise events from both data sources with 80 events that matched in both the app and the WHRM.

**Comparison of Exercise Behaviors between the App and WHRM**

The average weekly frequency of exercise was 3.0 (2.1) and 2.5 (1.8) days as tracked with the app and recorded by the WHRM, respectively \((p = 0.60)\). On average, participants who self-tracked their exercise with the app recorded an average of 0.3 (1.1) days more of exercise per week than the WHRM. The average duration from 80 matching exercise events from the app was 65.6 (55.2) minutes and from the WHRM was 64.8 (54.9) minutes \((p = 0.45)\), with a correlation of 0.90 \((\sim \text{Fig. 2})\).

Participants self-tracked moderate exercise in 46% (75/161) of the cases, with 32% (51/161) tracked as light, and 22% (35/161) as vigorous. The average heart rate of the WHRM exercise events were overwhelmingly categorized as moderate in 79% (92/116), with few categorized as vigorous or light, 18% (21/116) and 3% (3/116), respectively \((\sim \text{Fig. 3})\).

Participants self-tracked exercise intensity from the app matched the intensity as categorized by the average heart rate from the WHRM in 58% (46/80) and the highest heart rate in 40% (32/80) of the matched exercise events. When comparing against the average heart rate recorded by the WHRM during exercise, participants on average tracked a lower intensity in 35% (28/80) of exercise and a higher intensity in 7.5% (6/80). The Cohen’s kappa level of concordance for the app against the WHRM average heart rate was 0.2, a poor to fair level of agreement.

**Comparison of \( \Delta G/\text{min} \) between the App and WHRM**

There were 38 of the 277 exercise events for which there were no corresponding CGM readings, resulting in 239 exercise events where \( \Delta G/\text{min} \) could be calculated. The average \( \Delta G/\text{min} \) was greatest for light exercise self-tracked with the app, -0.53 (0.63) mg/(dL·min), while the average for vigorous exercise from the app was the least, -0.19 (0.89) mg/(dL·min). Moderate exercise from the app, light exercise from the WHRM, and moderate exercise from the WHRM were similar, -0.23 (0.99), -0.29 (0.48), and -0.29 (0.72) mg/(dL·min), respectively. Vigorous exercise from the WHRM actually had a slight increase of 0.09 (1.0) mg/(dL·min). The average \( \Delta G/\text{min} \) was not significantly different based on data source and intensity \((p = 0.21)\) and was found to be -0.27 (0.85) mg/(dL·min) \((\sim \text{Fig. 4})\).

**Discussion**

Measuring exercise behaviors for patients with diabetes is important to researchers, clinicians, and patients. Research studies often rely on exercise sessions for which the variables that influence \( \Delta G/\text{min} \) are highly controlled, while clinical settings often rely on retrospective self-report via questionnaires or surveys to assess physical activity or even more distal outcomes, such as HbA1c scores. Retrospective self-report through surveys and interviews, often the most convenient method for assessment, has shown limited validity and reliability at the individual level for providing insight that can be
In clinical settings, providers use electronic health records with varying levels of sophistication of clinical decision support for diabetes, with few systems capable of incorporating patient-generated data. Therefore, clinicians are faced with interpreting patient-reported data in the form of paper and digital logs as well as data generated from various devices, such as activity trackers, without any guidance.

Others have noted that technical errors occur with wearables devices, such as unreliable measurements and missing data, which we also experienced in this study. For example, one participant’s insulin pump was shifted for 1 hour, while another participant could not upload their WHRM data. Accelerometers, pedometers, multiple sensing systems, global positioning systems, and mathematical models are all methods that can be used to assess physical activity. While each method has its own strengths and weaknesses, for our study we chose to use an off-the-shelf WHRM with proprietary algorithms to gather information on specific exercise events rather than daily totals of activity. The testing of the validity...
The low- and high-end limitations of WHRM may have in concordance with the gold standard electrocardiogram. especially available heart rate monitors known to lose accuracy for and reliability of wearables is still underway, with commercially available heart rate monitors known to lose accuracy for very low and very high heart rates, but overall having high concordance with the gold standard electrocardiogram. The low- and high-end limitations of WHRM may have influenced the exercise events to primarily be categorized as moderate intensity in our study. Unfortunately, in our study there were technical difficulties that resulted in lost data, and possibly erroneous values due to user entry mistakes and inaccurate device measurements. For example, 80 events aligned between the app and the WHRM. Although this study was not designed to capture the reasons for unmatched data points, there were few cases where the WHRM detected physical activity not tracked by the participants. However, many of the unmatched cases could be due to not wearing, charging, or syncing the WHRM. The discrepancy could also be attributed to participants recording exercise with the app after the event and forgetting to adjust the app’s default time, which was the current time on the phone.

Diabetes devices, for example, insulin pumps, glucose meters, and CGM, are capable of storing objectively gathered data. The validity of CGM has been studied, with one recent study demonstrating that sensor accuracy of two commercially available CGM sensors was slightly diminished during physical activity when compared with a resting state. While we observed ΔG/min of −0.3 mg/(dL·min) from the CGM, a meta-analysis of ΔG/min in controlled exercise sessions found that for continuous moderate exercise ΔG/min was −1.33 mg/(dL·min), a fourfold larger impact on ΔG/min than what we observed. This discrepancy may be due to the controlled situations and differences in devices and methods for measuring glucose. For example, interstitial glucose measures, like those used in our study from the CGM, are known to lag behind blood glucose measurements, the common method used to validate CGM and the gold standard in exercise studies of patients with T1D. It may also be due to the participants not receiving specific instructions to calibrate the CGM and that sensors from different manufacturers were used by the participants, which would introduce noise into the glucose data. Also, participants in this study engaged in compensation techniques to offset the effects of exercise, such as disconnecting from the pump or taking a snack, as previously reported on this study cohort.

Our study is the first to take heterogeneous data from three sources used by patients in clinical settings and compare exercise behaviors to better understand their concordance in free-living patients. We had a small sample size that was mostly female, all of which owned smartphones, which limits the external validity of this study. Unique to this study and prior work by the authors is the focus on adult patients with T1D, as the majority of studies on exercise in free-living patients with T1D are conducted in youth and emerging adults. We found that participants self-tracking aligned well with the WHRM with respect to frequency and duration, which may have been due to the presence of the WHRM making patients more likely to self-track with the app. The intensity of exercise self-tracked had low levels of agreement with the WHRM. The discrepancy in intensity could be due to participants’ underreporting intensity, inaccuracies in WHRM due to limitations of the underlying technology to sense and calculate heart rate, and the thresholds chosen against the maximum heart rate could have misclassified exercise intensity. However, from a glucose management perspective, there was not a significant difference on ΔG/min.

One of the limitations of this study and others that have studied patient- and device-generated data is that there are no standard methods for analyzing and reporting the results. Most have focused on reporting overall ratios of desirable and undesirable behaviors related to logging glucose levels, while in this study we measured levels of agreement and presented averages from participants on behaviors related to exercise.

Our findings also indicate that self-tracking exercise behaviors with an app like iDECIDE, can be used to gather diabetes self-management behaviors and to facilitate decision making for self-management of glucose levels before engaging in exercise. This may be an attractive feature for researchers conducting studies and clinicians that provide care for patients with diabetes as it streamlines the processes for data gathering, assimilation, and analysis. As for patients on intensive insulin therapy who already must incorporate various supplies and devices into their daily regimen, it eliminates the need to incorporate an additional device to
wear, charge, and sync to interpret the data during and after exercise.

We found that self-tracked data from the app had very good agreement with the WHRM and that the effect on ΔG/min was constant across the devices and intensities. We are currently analyzing data that include survey and app data from this cohort and a recently recruited cohort to contrast survey responses about meal, alcohol, and exercise against behaviors self-tracked with the app. Future work includes continued research with the app in free-living settings to further study how best to incorporate patient-reported exercise into bedside tools to help clinicians better treat patients, such as providing data-driven personalized educational interventions.

**Conclusion**

We collated and analyzed data from three heterogeneous sources of data found in outpatient settings from free-living participants with T1D to answer clinically relevant questions. Clinicians and researchers should be aware of the various sources of patient-generated data and their levels of agreement when assessing exercise behaviors related to the self-care of T1D. Self-tracking exercise with an app and exercise recorded by a WHRM resulted in good levels of agreement and the same impact on ΔG/min. This result suggests that clinicians can rely on patient-reported exercise data when making treatment recommendations and that patients with T1D can make self-care decisions before engaging in physical activity based on their perceived intensity of the ensuing exercise instead of relying on a wearable device to assist with self-management of glucose levels during or after the physical activity.

**Clinical Relevance Statement**

iDECIDE, a smartphone app that allows patients to self-track exercise in real time, can reduce the number of devices patients incorporate into their self-care regimen by eliminating the need to use an additional activity monitoring device.

When making recommendations for incorporating exercise into self-management, clinicians can rely on exercise data gathered in real time, either self-reported with an app or recorded by a wearable device.

**Protection of Human and Animal Subjects**

This study was reviewed by the Arizona State University and Mayo Clinic Institutional Review Boards.

**Funding**

This research was supported by the 2018 ASU-Mayo Research Accelerator Award: Data-Driven Behavioral-Change Individualized Interventions to Improve Type 1 Diabetes.

**Conflict of Interest**

None declared.

**References**

7. Glasgow RE, Anderson RM. In diabetes care, moving from compliance to adherence is not enough. Something entirely different is needed. Diabetes Care 1999;22(12):2090–2092