ARTICLE

Continuous Monitoring Using a Wearable Device Detects Activity-Induced Heart Rate Changes After Administration of Amphetamine

Elena S. Izmailova[1†*], Ian L. McLean[1], Greg Hather[1], David Merberg[1] (iD) , Jason Homsy[1], Matthew Cantor[2], Dmitri Volfson[1], Gaurav Bhatia[2], Eric D. Perakslis[1], Christopher Benko[2] (iD) and John A. Wagner[1]

Wearable digital devices offer potential advantages over traditional methods for the collection of health-related information, including continuous collection of dense data while study subjects are ambulatory or in remote settings. We assessed the utility of collecting continuous actigraphy and cardiac monitoring by deploying two US Food and Drug Administration (FDA) 510(k)-cleared devices in a phase I clinical trial of a novel compound, which included the use of an amphetamine challenge. The Phillips Actiwatch Spectrum Pro (*Actiwatch*) was used to assess mobility and sleep. The Preventice BodyGuardian (*BodyGuardian*) was used for monitoring heart rate (HR) and respiratory rate (RR), via single-lead electrocardiogram (ECG) recordings, together with physical activity. We measured data collection rates, compared device readouts with conventional measures, and monitored changes in HR measures during the amphetamine challenge. Completeness of data collection was good for the Actiwatch (96%) and lower for the BodyGuardian (80%). A good correlation was observed between device and in-clinic measures for HR ($r = 0.99$; $P < 0.001$), but was poor for RR ($r = 0.39$; $P = 0.004$). Manual reviews of selected ECG strips corresponding to HR measures below, within, and above the normal range were consistent with BodyGuardian measurements. The BodyGuardian device detected clear HR responses after amphetamine administration while subjects were physically active, whereas conventional measures collected at predefined timepoints while subjects were resting and supine did not. Wearable digital technology shows promise for monitoring human subjects for physiologic changes and pharmacologic responses, although fit-for-purpose evaluation and validation continues to be important prior to the wider deployment of these devices.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
☑ Wearable digital devices have been adopted by many consumers. There is a limited number of published studies providing evidence that selected devices are appropriate for clinical trials.

WHAT QUESTION DID THIS STUDY ADDRESS?
☑ We assessed the utility of continuous actigraphy and cardiac data collection by wearable devices in the context of a phase I clinical trial, which included a novel compound in conjunction with activity-induced changes after amphetamine challenge.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
☑ We demonstrated acceptable data collection completeness for both devices, a good correlation between device and in-clinic measures for heart rate (HR) and less strong for respiratory rate. BodyGuardian was appropriate for continuous HR monitoring and detected clear changes in HR post–amphetamine challenge while subjects were physically active, whereas conventional measures collected at predefined timepoints did not.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?
☑ The current study suggests that utilization of wearable devices may provide advantages over traditional methods of vital sign data collection, particularly for detection of physiologic changes and pharmacologic responses.

Wearable digital technologies have been the subject of strong commercial promotion in recent years and have been adopted by many consumers. Although several devices have been approved by regulatory agencies and are increasingly used in healthcare settings, wearable devices have yet to find widespread application in industry-sponsored drug development studies.

Despite significant progress,[1–3] there are few published studies that include critical analysis of the technology at

---

[1]Takeda Pharmaceuticals International, Inc., Cambridge, Massachusetts, USA; [2]Koneksa Health Inc., New York, New York, USA. [†]Currently employed by Koneska Health, Inc. *Correspondence: Elena S. Izmailova (elena.izmailova@koneksahealth.com)

either the device or data-processing level. Some reports suggest that devices that collect vital sign (VS) data did not perform as the researchers had planned, requiring extensive manual review of data and time-consuming investigation of device-derived data artifacts.[4]

In early-stage drug development clinical trials, VS data are typically collected manually by clinical personnel: for example, measuring respiratory rate (RR) by timed direct observation of the subject's chest movements, or using clinic-based electronic devices to record data at discrete single timepoints, such as the use of a pulse oximeter or electrocardiogram (ECG) for measuring heart rate (HR). In the context of an industry-sponsored study following Good Clinical Practice (GCP) standards, these measurements are typically taken at a small number of predefined times: before, during, and after administration of the study drug while the subject is resident in a clinical pharmacology unit (CPU); during pre-exposure screening procedures; or during follow-up visits to the clinic. VS measurements are generally done following a period of rest (usually 5 minutes or longer, as defined in the study protocol or in the CPU's Standard Operating Procedures) in either the supine or sitting position. Additional assessments of VS data are performed in response to suspected safety or tolerability issues, or if the study drug or a challenge agent is expected to have pharmacological effects on VS.

The opportunity to record high-density VS data continuously using wearable digital sensors has the potential to (i) provide more information on study subjects' physiological profiles, and, therefore, offer greater sensitivity for detecting changes in these parameters; (ii) collect data during periods of physical activity; (iii) include periods of data collection in the subjects' normal home environment and during customary activities of daily living (ADL) rather than as inpatients in a residential CPU setting with protocol-imposed or practical constraints on activity level; and (iv) serve as an aid in interpretation of adverse events. In addition to providing more granular data and more complete detection of events, the use of this technology may reduce both the required duration of residential observation during phase I studies and the number of follow-up clinic visits.

Wearable digital devices may also improve evaluation of the impact of a novel medicine on disease activity or outcomes. In many therapeutic indications, the impacts of a drug on ADL and sleep patterns are important outcome measures. Changes in these parameters may indicate that the drug has had either a clinically relevant benefit—for example, an improvement in patients' mobility—or negative side effects, such as sleep disturbance. Currently, these assessments rely on patients' ability to recall these events in subsequent self-completed questionnaires. Such self-report is subjective and prone to confounding and recall bias[5] and may be improved by the inclusion of continuous real-time collection of activity-related data using digital devices to objectively monitor ADL.

We applied the principle of "fit for purpose" evaluation[6,7] to two wearable digital devices that have 510(k) device clearance from the US Food and Drug Administration (FDA): the Phillips Actiwatch Spectrum Pro (*Actiwatch*)[8]

and the Preventice BodyGuardian (*BodyGuardian*).[9] We incorporated the testing of these devices as an exploratory component of the TAK-041-1002 study,[10] a GCP single-site residential phase I study recruiting normal healthy volunteers to assess brain penetration of TAK-041 and its effects on amphetamine-induced dopamine release in the central nervous system (NCT number 02959892; EudraCT number 2016-002346-23; and Universal Trial Number: U1111-1184-1947)[10] using positron emission tomography (PET) imaging.

In a previous study of a wearable digital cardiac monitor,[11] we noted episodes of incomplete data collection and poor correlation between some VS parameters collected by the device and those collected by conventional in-clinic measures. We concluded that manual review of ECG strips from the device was necessary to interpret many of the signals recorded. In the current study, we (i) compared the wearable digital measures (using a different cardiac device) with the traditional VS data collected at the clinical site, (ii) examined HR modulation after amphetamine administration, and (iii) collected data concerning the completeness of data contribution by study subjects in the CPU and at home.

METHODS

The TAK-041-1002 clinical study was conducted in a residential CPU and a neuroimaging center, described elsewhere.[10] All subjects were healthy male volunteers recruited from the CPU's panel; they had no clinically significant acute or chronic medical disorders, were taking no concomitant medications, and had no exposure to other investigational or challenge agents in the 30 days preceding the study. The informed consent for the wearable component of the study was optional; 5 of 12 subjects consented to participate. The compound under development, TAK-041, penetrates into the human brain and has a target that is largely localized within the central nervous system. This study was conducted to assess the impact of TAK-041 on a neural pathway known to be activated by amphetamine challenge as assessed using a dopamine D2/D3 PET ligand. Expected effects of amphetamine in humans include increased physical activity and increased HR.[12]

The wearable digital devices were tested as an exploratory component of the protocol. Exploratory data were not included in the analysis of primary or secondary end points. Informed consent was obtained separately for the device component of the study, which was optional for any subject consenting to participate in the core part of the study. The study protocol was reviewed and approved by the relevant Ethics Committee (UK National Research Ethics Service (NRES) number 16/LO/1493). For the design of the study and authoring of the protocol and clinical study report documents, the VS, physical activity, and sleep data produced by wearable devices were treated as exploratory and were used for device evaluation purpose only. The data were not available to CPU or sponsor staff during the conduct of the study, and it was understood by the study execution team that these data were not intended to guide clinical care or other decision making during the conduct of the study.

Training in device use was provided to the staff of the CPU during in-person training, which included live demonstration of device application and data collection. Subjects were confined at a single residential CPU site in the United Kingdom for two study periods separated by a 5-day to 45-day interval (**Figure S1**). Clinical VS measurements were performed alongside collection of safety, tolerability, pharmacokinetic, and pharmacodynamic assessments.

Imaging was done at a nearby specialist imaging site, and included one magnetic resonance imaging scan of the head performed between screening and the first period of confinement as part of confirmation of study eligibility and to help delineate anatomic regions of interest for individual PET images and three PET scans preceded by intravenous administration of the radiolabeled dopamine D2/3 ligand [11]C-(+)-4-propyl-3,4,4a,5,6,10b-hexahydro-2H-naphtho[1,2-b][1,4]oxazin-9-ol (PHNO). During the first period of confinement, subjects had a baseline ^{11}C-PHNO PET scan. On the following day, each subject received amphetamine 0.5 mg/kg ~3 hours before the second PET scan. During the second period of confinement on day 1, the subjects received a single oral dose of TAK-041 followed by a single oral 0.5 mg/kg dose of amphetamine ~2 hours after the dose of TAK-041, followed by a [11]CPHNO PET scan at ~3 hours post–amphetamine administration. Both digital devices were applied between day −1 and day 1 of the first confinement period and on day 1 of the second confinement period (**Figure S1**), and were intended to be worn throughout the remainder of the confinement period in the CPU and during in-home monitoring for 5–6 days after discharge for the Actiwatch and 2 days after discharge for the BodyGuardian. Neither device was removed during PET imaging.

The Actiwatch[13] is worn on the wrist using a standard wristwatch-style strap and captures data on motion using a three-axis accelerometer; these data are used to derive information on activity level and sleep. The device has a 3-month battery life and a 30-day recording memory. Activity level is summarized using activity counts, a dimensionless measure of motion that is designed to remove the effects of gravity, transportation, and other acceleration that do not indicate subjects' physical activity. The BodyGuardian device[14] consists of an adhesive patch with skin electrodes and a sensor module for data collection, recording, and transmission of HR and RR biometric data. Following shaving of the skin (if necessary) and cleaning, the device is applied to the anterior surface of the left upper precordium using an adhesive strip. A duplicate device was applied approximately every 12 hours during charging of the first device, and adhesive strips were replaced as needed if loss of adhesion was apparent on visual inspection or was reported by the device (via an audible tone and screen message to the subject from the iPhone when connectivity dropped below 85%); typically, this was approximately every 48 hours. In-clinic HR was collected using the SpaceLab blood pressure monitor. In-clinic RR was collected manually by the site staff by observing the subjects' chest wall movements, counting respiration cycles over a defined time period, and entering this information immediately into the site's system, together with the time of data entry.

Wearable device data collection

The data collected by the Actiwatch were retrieved by periodically connecting it to a laptop computer running study-specific software, which downloaded the epoch level data from the device and saved them to a cloud-based database. The BodyGuardian device recorded a single-lead wall ECG via the two inner electrodes attached to the anterior chest. The electrode pads measure 10 mm diameter and have a signal sampling rate of 256 Hz with 12-bit resolution.[15] The device records ECG voltage every 8 ms, and from the resulting RR interval an estimate of HR was calculated approximately every 10 seconds, averaging six HR estimates per minute. The data collected by the BodyGuardian were streamed to a companion iPhone application (BodyGuardian Connect version 1.7.5) on a dedicated iPhone 5 via Bluetooth UHF radio technology. Data transmission occurred "live" when the iPhone was within Bluetooth range; this device had a recording memory of ~18 hours to allow for subsequent capture of data obtained when the device was unable to connect to the iPhone for real-time transmission. RR was estimated using the manufacturer's proprietary software algorithm from the cyclical fluctuation in HR associated with physiological sinus arrhythmia.

To estimate the completeness of data collection for each subject, invalid readings from the BodyGuardian were filtered (excluded) using the manufacturer's software during an initial quality control step: the recordings were sorted in timestamp order, and the durations of any gaps between valid recordings were determined. A similar approach was undertaken with the Actiwatch device: the periods when the device was not worn were filtered out (excluded). We also calculated the difference between the total time for which data were collected and the interval between the first application of the device and removal of the device at the end of the study. Data completeness was calculated as: 100% × (1 – (device noncovered time)/total study time).

We calculated completeness of data collection separately for each individual subject and each device. We used the standard millisecond coverage technique, which accounts for a variable epoch rate to estimate compliance. Our compliance estimate was the percentage of on-study milliseconds within a specified gap time (*T*) of a valid reading. For the BodyGuardian, we used an allowable gap time of *T* = 11 seconds. For the Actiwatch, we used an allowable gap time of *T* = 30 seconds. We did not attempt to confirm directly with the subjects whether longer periods of absent data were attributable to removal of the device by the subject rather than poor electrode contact or other "accidental" causes of loss of connection, nor did we question subjects regarding the reason(s) for unplanned removal of a device.

To calculate the number and duration of relevant time gaps of the BodyGuardian HR data (epoch = 10 seconds), data from each subject were sorted in timestamp order and the intervals between valid recordings were calculated. If an interval was >60 seconds, it was considered a gap. This interval length was selected as a minimal clinically relevant time when a safety signal could be missed. Total noncovered time was calculated by summing the length of all gaps.

Statistical analysis

Analysis was performed using R version 3.2.2 (R Foundation, Vienna, Austria). To compare the clinical and wearable measurements, the recorded collection times for the conventional clinical HR and RR measurements were first matched to the BodyGuardian epochs using a 1-minute interval. For example, if a conventional clinical HR measurement for a given subject was reported at 8:05 AM, then all BodyGuardian HR measurements for that subject from 8:04:30 AM to 8:05:30 AM were extracted. The mean of these BodyGuardian measurements within each interval were then compared with the corresponding clinical measurements. If no valid BodyGuardian data existed in a given time range, then the data point for that patient was excluded from analysis.

To compare the HR and RR measurements reported by the BodyGuardian device to the time-matched clinic measurements, the Pearson correlation between the wearable measurements and the clinical measurements was computed. Additionally, to assess the potential impact of treatment on the relationship between in-clinic and wearable measures, a mixed effect model was used. Briefly, the model assumed: BodyGuardian = (subject) + (treatment) + (clinical measurement) * (treatment), where the subject term was a random effect and the remaining terms were fixed effects. The amphetamine treatment condition was a binary factor designated "early" if the measurement time was <4 hours after amphetamine challenge and "late" otherwise.

Identification and characterization of L5 (5 hours of low activity) and M10 (10 hours of high activity) periods was as described elsewhere.[16] A Pearson correlation test was used to assess the strength of association between BodyGuardian and Actiwatch devices for L5 and M10 periods.

To quantify HR response to physical activity under the amphetamine challenge, we used two measures, the mean values of HR averaged over time, along with the fluctuations around the slow changes in HR averaged over time. These measures were derived from the signal obtained by the BodyGuardian device. They were calculated for the periods when subjects were resting and supine during the PET scan procedure (rest periods) and 1 hour immediately subsequent to the PET scan procedure excluding timepoints when subjects were resting and supine during the VS data collection procedure (active periods). In order to quantify fluctuations around the slow changes in HR, we used locally weighted scatterplot smoothing[17] with a span value of 0.5 to generate a smooth curve representing the trend in HR over a period of 1 hour (**Figure S6**). We then computed the median absolute deviation of the regression residuals. To test the difference of HR measures during the rest and active periods under the amphetamine challenge, we applied mixed model analysis of variance (ANOVA), with activity (active or rest) and period, and their interaction as fixed factor and subject as random factor. Significant ANOVA findings were followed by *post hoc* tests for least square means. All statistical tests were performed two-tailed at a 5% level of significance.

RESULTS

Patient demographics and study conduct

Five of the 12 subjects enrolled in the core clinical and imaging study consented to participate in the exploratory wearable digital device evaluation component. Reasons given for nonconsent included an expectation that compensation would be offered by the sponsor for additional study procedures. The participants were all men (per protocol), median age 47 years (range 34–55 years), and median body mass index 27.9 (range 26.0–29.1; **Table S1**).

One subject (57001-028) was discontinued due to failed PET ligand synthesis. The subject consented to participate in the wearable device component of the study during period 1; no data were collected for the study period 2. Subject 57001-017 consented to the wearable device component only for period 2 of the study (**Figure S1**). Conduct of the imaging study was otherwise uneventful.[10] There were no discontinuations attributable to the study drug, challenge agent, or other study procedures, no serious adverse events, and no significant new safety or tolerability findings attributable to TAK-041.

Completeness of data collection

Completeness of data collection for the Actiwatch was above 96% for all subjects during both the residential confinement and the at-home follow-up during both study periods (**Table S2**). For the BodyGuardian, data completeness ranged from 53–95% while the subjects were in the unit and from 69–96% while subjects were at home (**Table S2**). Periods of loss of valid data for the BodyGuardian device were attributed to intermittent poor skin contact and to subjects removing the devices for unspecified reasons. Additionally, the number of missing data gaps was variable across study subjects; the median gap length ranged from 90–120 seconds, and the longest gap varied from 1.7–102.7 hours (**Table S3**).

Comparison of in-clinic and wearable device measures

We performed comparison of conventional measurements made at the site with time-matched BodyGuardian measurements for HR and RR. Comparison of the paired HR data demonstrated a strong correlation between in-clinic and wearable device measurements (Pearson's $r = 0.99$; $P < 0.001$; **Figure 1**). The RR measures derived from the BodyGuardian device and corresponding in-clinic measures were poorly correlated with a Pearson's correlation coefficient of 0.39 ($P = 0.004$). We used a mixed effect model to explore the relationship between the in-clinic and mobile data adjusting for time and treatment effects. The mixed effect model using a mobile device HR explained 98% of the variation in the in-clinic counterpart, whereas the model using BodyGuardian RR data explained only 25% of the variation in the in-clinic measurement.

Additionally, we assessed the 95% limits of agreement between BodyGuardian HR measures and conventional measures performed at the site. The 95% limits of agreement based on the Bland-Altman method were –3.2 to 5.7 bpm. This range constitutes 14% of the mean conventional HR
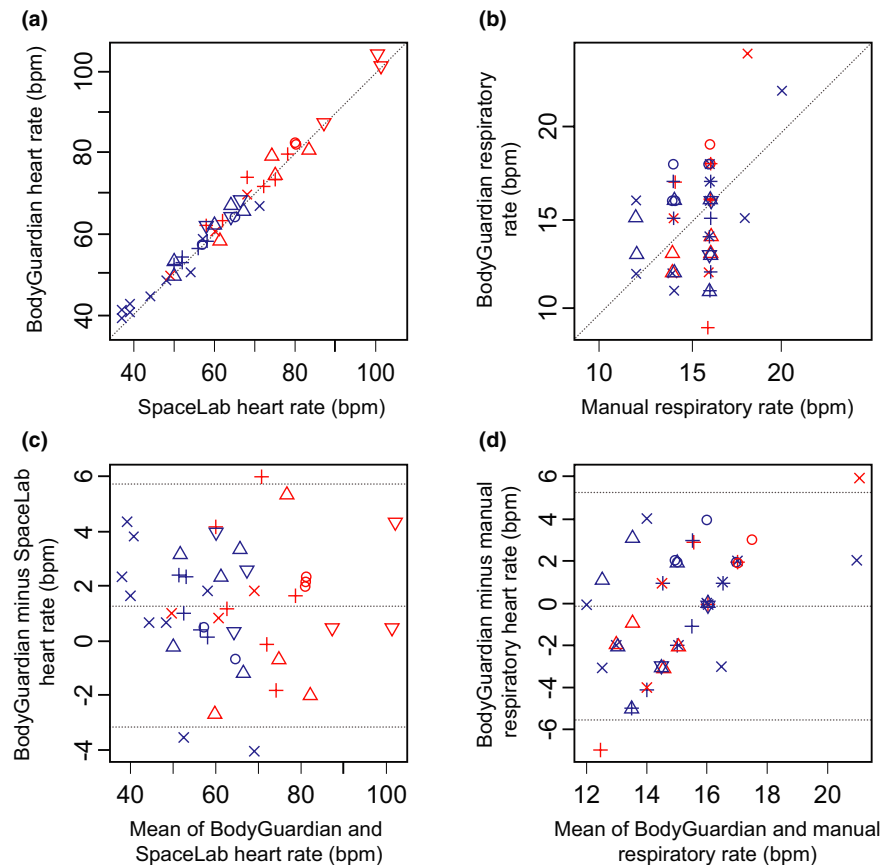
Figure 1 The comparison of in-clinic and wearable measurements for heart rate and respiratory rate. (**a,b**) Correlation and (**c,d**) Bland Altman analyses for in-clinic and wearable device measures for (**a,c**) heart rate and (**b,d**) respiratory rate. In-clinic heart rate measurements were performed using the SpaceLab device, in-clinic respiratory rate measurements were performed using the manual method; the BodyGuardian Device was used for wearable heart rate and respiratory rate measurements. Different symbols depict individual study subjects; blue color indicates time points taken prior to amphetamine challenge, red color indicates time points post-amphetamine challenge. For heart rate bpm stands for beats per minute; for respitory rate bpm stands for breaths per minute.

(**Figure 1**, Bland-Altman plot). Overall, the wearable device HR data were highly consistent with the in-clinic counterpart.

The Bland-Altman limits of agreement indicated poor agreement between RR in-clinic measures and their counterparts derived from the wearable device. The 95% limits of agreement were –5.5 to 5.2 breaths/minute, corresponding to 71% of the mean RR. Overall, the wearable device RR measurements did not show strong agreement with the in-clinic measurements.

Comparison of movement assessment between devices

We used two separate accelerometers to measure study subjects' physical activity to facilitate the interpretation of VS: one body-worn (BodyGuardian device) and one wrist-worn (Actiwatch device). In order to assess concordance between the two accelerometers, we measured physical activity during periods of rest and high physical activity. We used the approach of identifying L5 (low physical activity) and M10 (high physical activity) periods for each study subject as described elsewhere.[16] Given the typically high level of variation observed in subjects' activity during 24-hour intervals, separating correlation between two different accelerometers in two periods of light and high activity provided

more refined data concerning the utility of each. Comparing physical activity during selected L5 periods indicated that both accelerometers consistently detected periods of low physical activity. However, the correlation was poor, ranging from $r – 0.05$ to $r = 0.52$ (the latter driven largely by one timepoint); the range of values was low (**Figure S3**). The physical activity data from selected M10 periods were better correlated than the data from the L5 period, with correlation coefficients ranging from $r = 0.3$ to $r = 0.84$ and exhibiting substantial variation among subjects and days (**Figure S4**). Again, the data from both accelerometers were useful to detect periods of high physical activity. However, the correlation was highly variable across subjects. Additionally, we identified a need to discriminate time intervals when a device is worn by a subject and not to confuse periods of inactivity with time intervals when the device was not worn with a charged battery. The Actiwatch device has an output of "device worn/not worn," which makes it possible to filter out the periods when the device is not worn. The BodyGuardian device does not have a similar function. The periods when the device was not worn were excluded after a manual data review by checking for presence of VS data (**Figure S5**).

Face validity of VS and sleep data

The HR and RR data were distributed as anticipated; conforming to the ranges typical for the normal healthy volunteer population recruited for a phase I study (HR ~50–120 bpm, RR ~10–16 breaths/minute). To estimate the proportion of measures reported by the BodyGuardian device outside of the normal range, we calculated the percent of epochs with HR values <50 bpm and >120 bpm. The percent of epochs within the normal range were ranging from 92–98% in four of five study subjects. Subject 57001-025 had 19.91% of epochs with values below 50 bpm (**Table 1**). Further examination of these subjects' data showed that most of these measures were above 45 bpm (**Table S4**). Manual calculations of HRs by measuring the R-R intervals on selected ECG strips at HR values below, within, and above the normal HR range were consistent with the respective values reported by the BodyGuardian device (**Figure 2**).

We also examined measures derived from the Actiwatch mobility data, such as sleep parameters, including total sleep time (TST) and sleep efficiency (SE). The data analysis suggested that sleep time ranged from 4.6–8.3 hours, with substantial variation among subjects, and the mean SE was >80% in four of five subjects (**Table S5**).

Amphetamine challenge

The study included the amphetamine challenge in both presence (period 2) and absence (period 1) of the experimental study drug. Using wearable device data, we observed an increase of HR following the amphetamine challenge similar in magnitude, as described earlier.[18] Additionally, we observed markedly increased HR when study subjects were ambulatory and physically active, as detected by both BodyGuardian and Actiwatch accelerometers. A representative example of such HR increase is shown in **Figure 3**. Importantly, HR measures were lower when subjects were resting and supine during the periods of conventional VS data collection, according to the study protocol, and did not change substanatially post–amphetamine challenge (**Figure 3**). Moreover, we calculated an average HR post–amphetamine challenge during both period 1 and period 2 for each subject for the duration of the PET scans (resting and supine) and 1 hour immediately after the PET scans (active). ANOVA models for the mean HR values and for the amplitude of fluctuations in HR revealed no evidence for interaction effect between period and activity; therefore, only linear terms were retained. In the case of mean HR, both activity and period were significant factors ($F(1,8) = 36.8$; $P = 0.0003$ and $F(1,8) = 7.0$; $P = 0.03$, respectively), and HR was

Table 1 Percent of epochs with HR value below, within, and above normal range

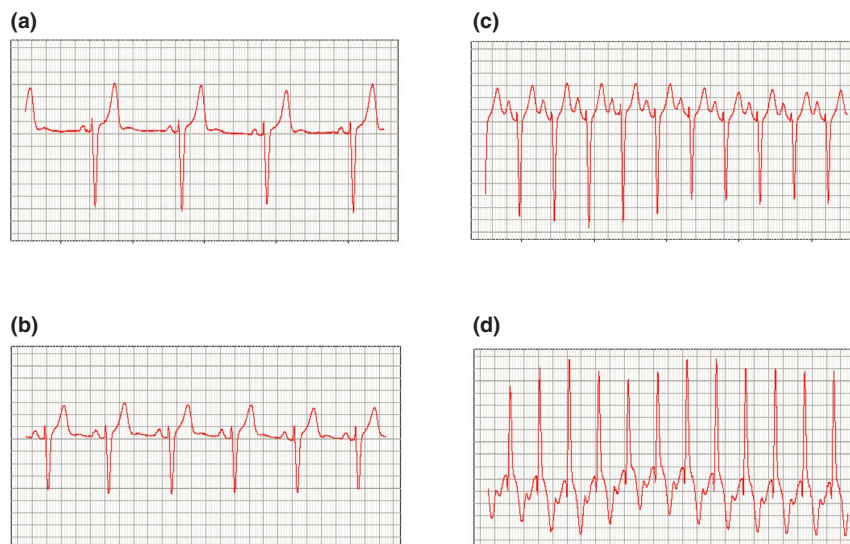| Subject ID	% of epochs HR ≥0 and <50 bpm	% of epochs HR ≥50 and <120 bpm	% of epochs HR ≥120 and <150 bpm	% of epochs HR ≥150 and <180 bpm	% of epochs HR ≥180 bpm
57001-017	0.01	98.47	1.52	0	0
57001-022	0.25	96.53	3.2	0.03	0
57001-024	0.01	92.26	7.72	0.01	0
57001-025	19.91	76.55	3.51	0.03	0
57001-028	0	96.95	3.04	0	0

bpm, beats per minute; HR, heart rate.

Figure 2 Examples of electrocardiogram strips corresponding to (**a**) heart rate (HR) = 49 bpm, (**b**) HR = 63 bpm, (**c**) HR = 125 bpm, and (**d**) HR = 151 bpm. The x-axis scale is 200 ms per the major gridline; the y-axis scale is 0.2 mV per the major gridline. bpm, beats per minute.
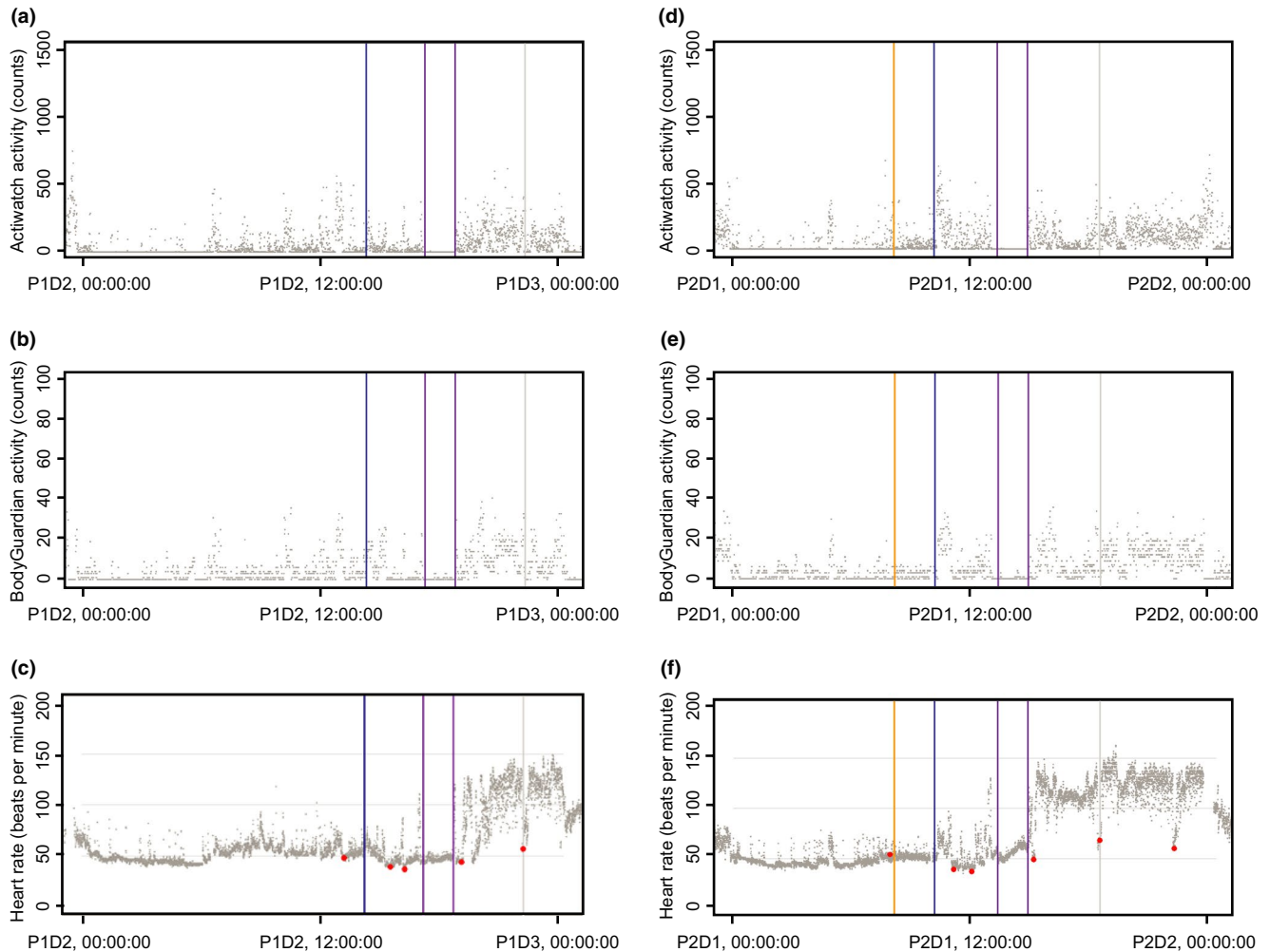
Figure 3 Changes in heart rate (HR) before and after amphetamine challenge in (**a–c**) period 1 and (**d–f**) period 2 of the study. Gray dots indicate data produced by wearable devices: activity counts by Actiwatch (**a, d**), activity counts by the BodyGuardian device (**b, e**), and HR (**c, f**). Red dots indicate HR measurements done in clinic by the conventional method. The blue line indicates the timing of amphetamine challenge, the magenta line indicates the start and end of positron emission tomography scans, and the yellow line indicates the time of TAK-041 dose during period 2. D, day; P, period; P1D2, period 1 day 2.

significantly higher in active state ($t(8) = 6.1$; $P = 0.0003$; **Figure 4a**). In the case of the amplitude of fluctuations, activity was a significant factor ($F(1,8) = 49.5$; $P = 0.0001$), and level of fluctuations was significantly higher in the active state compared with the next hour ($t(8) = 7.6$; $P = 0.0001$; **Figures 4b and S6**).

DISCUSSION

This study assessed the utility of continuous monitoring of VS in the context of drug development. We used two 510(k)-cleared wearable digital devices to collect HR and RR data, as well as activity data that we used to interpret ambulatory VS data. The device study was embedded in a residential phase I therapeutic study conducted under GCP. We performed additional validation experiments of the above-mentioned devices to confirm that the devices were adequate for use in human experimentation. The purpose of the FDA 510(k) clearance program is to provide a path forward to

legally market medical devices with demonstrated equivalency to a predicate device. However, using these devices in the context of clinical trials according to an intended use, as it was in our study, presents an additional layer of challenge because of requirements such as data reporting to the regulatory agencies.

The main goal of the wearable device study was to collect dense HR data in the context of amphetamine administration during periods of rest and physical activity as a model for potential HR safety signals. Additional goals included comparison of wearable device measures to conventional VS readouts and collection of patient compliance data.

The BodyGuardian and Actiwatch devices did not interfere with the other study procedures, including dosing of the study drug or challenge agent, collection of conventional VS data, safety and pharmacokinetic sample collection, or multiple PET scans, and were well tolerated by consenting study subjects. Our results indicate that deploying chest-worn and wrist-worn wearable devices in a phase I study of
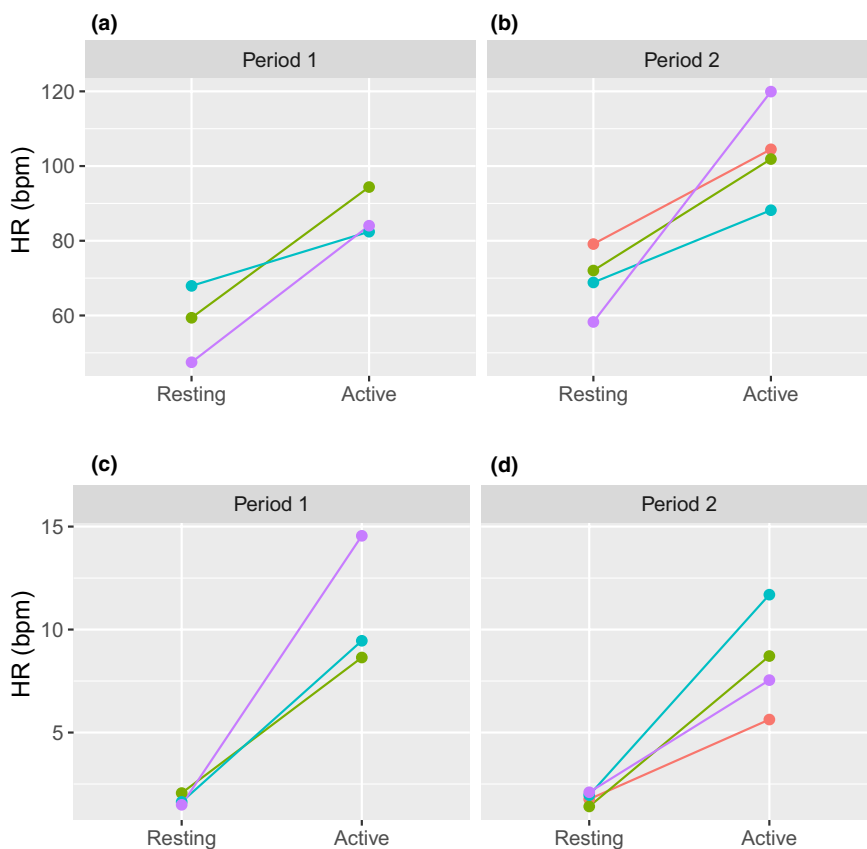
Figure 4 BodyGuardian heart rate (HR) after amphetamine challenge during rest and active periods. (**a, b**) Mean HR during periods. (**c, d**) Mean absolute deviation of residuals from smoothed curve. Colors indicate individual subject data.

relatively high complexity is feasible, resulting in acceptable compliance (**Table S2**) and yielding sufficient data to interpret study results.

Inspection of wearable device VS data indicated that the distribution of HR values was as expected in ambulatory subjects (**Table 1**). The sleep data were also consistent with previously reported results.[11] TST ranged from 6.5–8.3 hours and SE was >80% in four of five subjects (**Table S5**). These values are smaller than typical sleep times for healthy adults and may be a consequence of sleeping in an unfamiliar hospital environment. TST and SE were lowest in subject 57001-028, but this individual contributed data for only 2 days during period 1 (**Figure S2**). This time period may be insufficient to establish reliable TST and SE parameters by means of actigraphy.[19] We also observed data gaps, due to devices not being worn or data not meeting acceptance criteria prespecified by the device manufacturers. This is an issue inherent to continuous monitoring as described earlier.[1,4] An important objective of studies like this one, however, is to establish expectations for both subject compliance in contributing the data, as well as the rate of available valid data in order to enable realistic expectations of data availability in future studies.

We found that the device-reported HR values agreed with manual inspection, suggesting the technology is appropriate for continuous HR monitoring. The advantage of using the BodyGuardian device compared with traditional monitoring devices, such as Holter monitors, is the option of near real-time monitoring of data (although this feature was not used in the present study).

We also compared HR and RR measures to the conventional methods of conventional VS data collection performed by the site. Similar to our previous study,[11] HR data showed a strong correlation and tight limits of agreement with the traditional counterpart. RR data demonstrated a weak correlation and wider limits of agreement, as described previously,[11] which likely reflects the difference between manual and device-based methods of data collection.

Additionally, we demonstrated the importance of monitoring subjects' physical activity to interpretation of the VS data. Our comparison of mobility counts between Actiwatch and BodyGuardian devices indicated that both can reliably distinguish periods of high and low physical activity accurately, as demonstrated by concordance between L5 and M10 periods predicted by both devices. The correlation for any given period was variable in ways that could be explained by device placement, and also may be impacted by physical activity type. We found the Actiwatch device to be more convenient, as it automatically provided information on when the device was or was not worn, allowing quick elimination of invalid data; in contrast, using the accelerometer data from the BodyGuardian device required an additional examination of the HR data to gain the same understanding.

Finally, we examined HR data collected following amphetamine challenge and detected marked increases in HR only

when subjects were physically active (**Figure 4**). If commonly used methods of VS data collection (resting and supine for a defined period) were performed, the HR increases occurring between traditional VS data collection timepoints would have been missed (**Figure 3**). Therefore, our data illustrate the power of continuous data monitoring for detection of transient physiological changes and pharmacological responses. Such data can be informative in early clinical development to guide decisions about investigational drugs as well as detecting and understanding safety signals earlier in the development life cycle.

Our study has certain limitations. The small study size requires confirmation of these findings in a larger clinical trial. We used the amphetamine administration component of the study as a model setting to demonstrate the utility of continuous VS monitoring using previously published results.[12,18] However, detecting unanticipated safety signals represents a significantly greater challenge, as it will require more careful examination of the findings to rule out potential artifacts. Although Holter ECG monitoring has been used for continuous data collection and clinical trials, this method is deployed only for limited periods of time and the data are available only after the completion of the data collection. Arrhythmia detection via a single lead ECG device has convenience advantages relative to conventional Holter monitoring, as has been described.[3,20,21] Additionally, a comparison of the BodyGuardian device data to an independent device data, such as Holter, would help to facilitate the decision-making process around the choice of the device. However, the requirement for manual review shared by both BodyGuardian and Holter devices to confirm anomalous cardiovascular signals detected during continuous monitoring[4] is an important consideration.

In conclusion, we found that continuous monitoring of HR advantageous for detecting certain cardiovascular events, such as tachycardia, which sparsely applied conventional methods may miss. However, the decision to deploy continuous monitoring needs to be taken carefully, as it currently requires manual review of potentially aberrant signals by a qualified professional,[22] which can be resource intensive. Recent developments in the area of wearable devices, both consumer and medical grade, provide opportunities to transform drug development by enabling new, more patient-centric and cost-efficient approaches to data collection. However, decisions about device selection and mode of use should be based on scientific evidence. The present study provides such evidence, indicating that utilization of wearable devices may provide advantages over traditional methods of VS data collection.

Supporting Information. Supplementary information accompanies this paper on the *Clinical and Translational Science* website (www.cts-journal.com).

Author Contributions. E.S.I. and I.L.M. wrote the manuscript. E.S.I., I.L.M., M.C., G.B., E.D.P., C.B., and J.A.W. designed the research. E.S.I., I.L.M., J.H., M.C., G.B., C.B., and J.A.W. performed the research. E.S.I., G.H., D.M., J.H., G.B., and D.V. analyzed the data.

1. Izmailova, E.S., Wagner, J.A. & Perakslis, E.D. Wearable devices in clinical trials: hype and hypothesis. *Clin. Pharmacol. Ther.* **104**, 42–52 (2017).
2. Steinhubl, S.R., Kim, K.I., Ajayi, T. & Topol, E.J. Virtual care for improved global health. *Lancet* **391**, 419 (2018).
3. Snyder, C.W., Dorsey, E.R. & Atreja, A. The best digital biomarkers papers of 2017. *Digit. Biomark.* **2**, 64–73 (2018).
4. Weenk, M. *et al.* Continuous monitoring of vital signs using wearable devices on the general ward: pilot study. *JMIR Mhealth Uhealth* **5**, e91 (2017).
5. Schmier, J.K. & Halpern, M.T. Patient recall and recall bias of health state and health status. *Expert Rev. Pharmacoecon. Outcomes Res.* **4**, 159–163 (2004).
6. Lee, J.W. *et al.* Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm. Res.* **23**, 312–328 (2006).
7. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, & other Tools) Resource. (US Food and Drug Administration, Silver Spring, MD, 2016).
8. Actiwatch 510(k) clearance. <https://www.accessdata.fda.gov/cdrh_docs/pdf/K983533.pdf>. Accessed April 11, 2019.
9. BodyGuardian 510(k) clearance. <https://www.accessdata.fda.gov/cdrh_docs/pdf15/K151188.pdf>. Accessed April 11, 2019.
10. A phase 1, open-label, positron emission tomography study in healthy subjects to determine the effect of TAK-041 on amphetamine-Induced dopamine release in the CNS after single-dose oral administration. <https://clinicaltrials.gov/ct2/show/NCT02959892>. Accessed April 11, 2019.
11. Izmailova, E.S. *et al.* Evaluation of wearable digital devices in a phase I clinical trial. *Clin. Transl. Sci.* **12**, 247–256 (2019).
12. Martin, W.R., Sloan, J.W., Sapira, J.D. & Jasinski, D.R. Physiologic, subjective, and behavioral effects of amphetamine, methamphetamine, ephedrine, phenmetrazine, and methylphenidate in man. *Clin. Pharmacol. Ther.* **12**, 245–258 (1971).
13. Actiwatch Spectrum Activity monitor. <http://www.usa.philips.com/healthcare/product/HC1046964/actiwatch-spectrum-activity-monitor>. Accessed April 11, 2019.
14. BodyGuardian device. <http://preventicesolutions.com/services/body-guardian-heart.html>. Accessed April 11, 2019.
15. Bruce, C.J. *et al.* Remote electrocardiograph monitoring using a novel adhesive strip sensor: a pilot study. *World J. Cardiol.* **8**, 559–565 (2016).
16. Jones, S.E. *et al.* Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. *Nat. Commun.* **10**, 1585 (2019).
17. Wang, Y., Han, F., Zhu, L., Deussen, O. & Chen, B. Line graph or scatter plot? Automatic selection of methods for visualizing trends in time series. *IEEE Trans. Visual Comput. Graphics* **24**, 1141–1154 (2018).

686

18. Weidner, L.D., Paris, A., Frankle, W.G. & Narendran, R. Safety of oral mmphetamine administered during positron emission tomography scans in medically screened humans. *PLoS One* **10**, e0140647 (2015).

19. Aili, K., Astrom-Paulsson, S., Stoetzer, U., Svartengren, M. & Hillert, L. Reliability of actigraphy and subjective sleep measurements in adults: the design of sleep assessments. *J. Clin. Sleep Med.* **13**, 39–47 (2017).

20. Barrett, P.M. *et al.* Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring. *Am. J. Med.* **127**, 95.e11–7 (2014).

21. Cheung, C.C., Kerr, C.R. & Krahn, A.D. Comparing 14-day adhesive patch with 24-h Holter monitoring. *Future Cardiol.* **10**, 319–322 (2014).

22. Zimetbaum, P. & Goldman, A. Ambulatory arrhythmia monitoring: choosing the right device. *Circulation* **122**, 1629–1636 (2010).