



Validation of accelerometry as a digital phenotyping measure of negative symptoms in schizophrenia

Gregory P. Strauss¹✉, Ian M. Raugh¹, Luyu Zhang¹, Lauren Luther¹, Hannah C. Chapman¹, Daniel N. Allen², Brian Kirkpatrick³ and Alex S. Cohen⁴

Negative symptoms are commonly assessed via clinical rating scales; however, these measures have several inherent limitations that impact validity and utility for their use in clinical trials. Objective digital phenotyping measures that overcome some of these limitations are now available. The current study evaluated the validity of accelerometry (ACL), a passive digital phenotyping method that involves collecting data on the presence, vigor, and variability of movement. Outpatients with schizophrenia (SZ: $n = 50$) and demographically matched healthy controls (CN: $n = 70$) had ACL continuously recorded from a smartphone and smartband for 6 days. Active digital phenotyping assessments, including surveys related to activity context, were also collected via 8 daily surveys throughout the 6 day period. SZ participants had lower scores on phone ACL variables reflecting vigor and variability of movement compared to CN. ACL variables demonstrated convergent validity as indicated by significant correlations with active digital phenotyping self-reports of time spent in goal-directed activities and clinical ratings of negative symptoms. The discriminant validity of ACL was demonstrated by low correlations with clinical rating scale measures of positive, disorganized, and total symptoms. Collectively, findings suggest that ACL is a valid objective measure of negative symptoms that may complement traditional approaches to assessing the construct using clinical rating scales.

Schizophrenia (2022)8:37; <https://doi.org/10.1038/s41537-022-00241-z>

INTRODUCTION

Negative symptoms significantly limit potential for successful functional outcome and recovery in individuals with schizophrenia^{1–3}. Unfortunately, currently available psychosocial and pharmacological treatments yield minimal benefits for negative symptoms⁴ and no drug has received United States Food and Drug Administration (FDA) approval for an indication of negative symptoms.

In an effort to make progress in the treatment of negative symptoms, the United States National Institute of Mental Health (NIMH) hosted a consensus development conference in 2005. Several key conclusions resulted from this meeting. Among these were that: 1) at least 5 core domains exist within the negative symptom construct (blunted affect, avolition, anhedonia, asociality); and 2) first-generation clinical rating scales (e.g., SANS, BPRS, PANSS, NSA) were inadequate for measuring the construct, and new assessment tools were needed to evaluate the 5 consensus domains according to modern conceptualizations⁵.

Following the Consensus Development conference, a workgroup was formed to create a clinical rating scale that would meet these needs in the field. In the process, it became clear that two instruments were needed: one that would address the essential elements of each domain but be concise enough to be practical for routine clinical use as well as large multicenter clinical trials and another that would cover each of the domains in extensive detail. As a consequence, the workgroup divided into 2 groups to accomplish these goals and two second-generation rating instruments were developed: the Brief Negative Symptom Scale (BNSS)⁶ and Clinical Assessment Interview for Negative Symptoms (CAINS)⁷. Excellent

psychometric properties have been demonstrated for the original English versions of the BNSS^{6,8–13} and the CAINS^{7,14,15}, as well as translated versions of both scales^{16–32}. Although there is considerable overlap in how the BNSS and CAINS assess the 5 domains identified in the consensus conference, they each have important practical and conceptual differences that offer distinct advantages for use in clinical trials and laboratory-based experimental psychopathology studies¹¹.

Although second-generation scales, such as the BNSS and CAINS, represent an important advance from first-generation scales, they are subject to several limitations that are inherent to all clinical interview-based measures. For example, clinical ratings are: (1) influenced by cognitive impairments (e.g., long-term and working memory impairments) that make retrospective and prospective reports less accurate³³; (2) subject to biases and self-report confounds, such as social desirability, over and under reporting tendencies, halo effects, and biases resulting from patient/interviewer characteristics (e.g., gender, ethnicity); (3) impacted by limited precision and resolution (e.g., requiring raters to average across distinct contexts and lengthy time intervals that may hold meaningful variance). Clinical rating scales are also costly in terms of financial and time costs needed to implement them in clinical trials, and they may not be highly sensitive to treatment effects.

It may now be possible to circumvent these limitations associated with clinical rating scales using what may become the “third generation” of negative symptom assessment: digital phenotyping. Digital phenotyping refers to the use of mobile devices (e.g., wearable smartbands, smart phones) to initiate the collection of data in everyday life^{34,35}. A distinction is made between digital phenotyping approaches that are passive versus

¹Department of Psychology, University of Georgia, Athens, GA, USA. ²Department of Psychology, University of Nevada, Las Vegas, NV, USA. ³Department of Psychiatry and Behavioral Sciences, University of Nevada, Reno School of Medicine, Las Vegas, NV, USA. ⁴Department of Psychology, Louisiana State University, Baton Rouge, LA, USA.

✉email: gstrauss@uga.edu

active. Passive digital phenotyping involves the unobtrusive collection of objective data, typically from the internal sensors of a device (e.g., smartphone or band). Several types of passive digital phenotyping variables can be collected that may hold relevance to negative symptoms, such as phone usage data (e.g., call/text logs, screen time, Bluetooth connectivity), social media data (e.g., time spent on apps like Facebook or Instagram), geolocation (GPS coordinate data), speech samples collected from ambient sound in the environment, and ambulatory psychophysiology. These objective, passive digital phenotyping variables can be paired with active digital phenotyping methods that require the participant to choose to perform an activity or behavior of interest (e.g., a survey, video, or cognitive task on the phone). When passive data is scaled and epoched to the same time interval as the active data, the combined approach allows for a more nuanced examination of symptoms as they unfold in real life (e.g., when the participant is in a certain mood state, engaged in a social interaction, or in a certain location)³⁶. To date, few studies have evaluated the utility of combining active and passive digital phenotyping measures to evaluate negative symptoms in SZ.

The aim of the current study was to provide an initial psychometric evaluation of one promising passive digital phenotyping measure of negative symptoms: accelerometry (ACL). ACL involves calculating the presence, vigor, and variability of motor movement. ACL is similar to other measures of motor behavior in real life, such as actigraphy; however, unlike these other measures, which rely on specialized sensors/instruments designed for research purposes, ACL can be collected via the internal sensors of smartphones or smartbands that are commercially available and thus more feasible for routine use. Accelerometers have received approval as medical devices from the U.S. Food and Drug Administration for a number of uses related to daily activity, falls, and sleep. They have also been used as outcome measures in clinical trials measuring physical and mental health outcomes³⁷. When paired with active digital phenotyping approaches (e.g., surveys), ACL has the potential to lead to a more granular understanding of human activity across contexts (e.g., locations, activity types, social settings), including an understanding of goal-directed behaviors encapsulated in the negative symptom construct.

Few studies have explicitly evaluated accelerometry as a method for assessing symptoms of SZ. In an epidemiological study in the United Kingdom, Firth et al.³⁸ evaluated the utility of ACL for objectively evaluating physical activity relative to self-reported physical activity; although individuals with SZ had no self-reported reductions in physical activity, ACL indicated an objective reduction in amount of movement. Ben-Zeev and colleagues have used ACL in their Cross-Check symptom prediction and relapse monitoring platform, demonstrating that various ACL measures predict relapse and symptom severity on a 7-item version of the Brief Psychiatric Rating Scale^{39–41}; however, given that the BPRS does not measure the volitional component of negative symptoms⁴² and that convergence with aggregate BPRS scores was evaluated, it is unclear whether ACL metrics are valid measures of negative symptoms specifically.

To our knowledge, no study has directly explored the validity of ACL as a measure of negative symptoms in the context of everyday life, while pairing ACL with active digital phenotyping data to evaluate validity. We asked participants to perform 6 days of data collection, where active surveys evaluating symptoms and context (location, activity, social interaction) were sent 8 times per day and ACL data was collected continuously via smartphone and a wrist-worn smartband. The following hypotheses were made: (1) SZ would have lower scores than CN on ACL mean and standard deviation variables measuring the presence/vigor of movement and variability of movement; (2) convergent validity of ACL would be demonstrated with: concurrently collected active self-reports of goal-directed activity context, as well as clinically

rated negative symptoms measured via standard clinical interviews; (3) discriminant validity would be supported by low and nonsignificant correlations with measures of positive, disorganized, and total psychiatric symptoms collected via standard clinical rating scale measures.

RESULTS

Hypothesis 1: Group differences in accelerometry variables

SZ had significantly lower scores than CN for phone ACL mean, $F(1, 118) = 17.75, p < 0.001$, and ACL SD, $F(1, 118) = 23.64, p < 0.001$. Groups did not significantly differ on band ACL mean, $F(1, 64) = 0.06, p = 0.80$, ACL SD, $F(1, 64) = 0.60, p = 0.44$, or ACL Activity Index, $F(1, 64) = 0.47, p = 0.50$. Group comparisons are presented as Z-scores in Fig. 1 to facilitate ease of comparison across different units of measurement.

Exploratory analyses were also conducted for each of the phone and band ACL variables based on two types of contexts that should in theory demonstrate differences when participants are in that context versus not: public location context and recreational activity context (see Supplementary Table S1 in Supplemental Materials). Phone ACL showed significant or trend level higher mean and SD when participants in both groups reported being in public and while engaged in a recreational activity compared to when they were not in these contexts. Band measures did not demonstrate differences across contexts. These findings support the validity of the phone ACL metrics by indicating greater movement during real-world contexts where greater magnitude and variability of movement would be expected.

Hypothesis 2: Convergent validity

Convergent validity was demonstrated within the SZ sample via significant negative correlations between band accelerometry mean and SD measures and clinically rated negative symptoms on the BNSS. Phone ACL metrics were not significantly associated with negative symptoms on the BNSS (see Table 1).

The EMA data also supported convergent validity of ACL. The proportion of time spent in goal-directed activities was significantly correlated with phone ACL mean ($r = 0.45, p < 0.001$) and SD ($r = 0.39, p < 0.001$) variables, such that greater vigor and variability in movement were associated with more goal-directed activity across the entire sample of SZ and CN participants. Band ACL metrics were not significantly correlated with EMA reported goal-directed activity (r 's < 0.09).

We also explored the association between band and phone ACL scores. No associations were significant in schizophrenia (r 's < 0.23) or controls (r 's < 0.10) (see Supplementary Table S2 in Supplemental Materials for complete correlations).

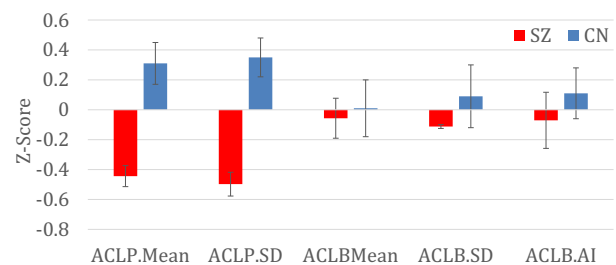


Fig. 1 Group comparisons for accelerometry measured via smartphone and smartband. SZ schizophrenia, CN control, ACLP.Mean accelerometry phone mean, ACLP.SD accelerometry phone average standard deviation, ACLB.Mean accelerometry band mean, ACLB.AI accelerometry band activity index. Error bars represent standard errors.

Table 1. Convergent validity correlations.

	ACL.P. mean	ACL.P.SD	ACL.B. Mean	ACL.B.SD	ACL.B.AI
BNSS anhedonia	-0.12	-0.21	-0.45*	0.28	0.33
BNSS asociality	0.09	-0.09	-0.34	0.11	0.33
BNSS avolition	0.16	-0.14	-0.41*	0.39*	0.38
BNSS blunted affect	-0.13	-0.18	-0.40*	0.36	0.27
BNSS alogia	0.17	0.02	-0.40*	0.27	0.26

ACL.P.Mean accelerometry phone mean, ACL.P.SD accelerometry phone average standard deviation, ACL.B.Mean accelerometry band mean, ACL.B.SD accelerometry phone average standard deviation, ACL.B.AI accelerometry band activity index, BNSS Brief Negative Symptom Scale.

Bold values indicate statistical significance $p < 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 2. Discriminant validity correlations.

	ACL.P. mean	ACL.P.SD	ACL.B. Mean	ACL.B.SD	ACL.B.AI
PANSS positive	0.25	0.14	0.00	0.12	0.01
PANSS disorganized	-0.01	0.24	-0.08	0.17	-0.17
PANSS total	0.12	0.27	-0.24	0.36	0.22

ACL.P.Mean accelerometry phone mean, ACL.P.SD accelerometry phone average standard deviation, ACL.B.Mean accelerometry band mean, ACL.B.AI accelerometry band activity index, PANSS Positive and Negative Syndrome Scale.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Hypothesis 3: Discriminant validity

Discriminant validity was supported by low and nonsignificant correlations between ACL phone and band measures and clinically rated positive, disorganized, and total symptoms on the PANSS (see Table 2). Additionally, the higher magnitude nonsignificant correlations were in the opposite direction of what would be expected to indicate poor discriminant validity (i.e., greater positive and general symptoms associated with greater movement).

DISCUSSION

The current study evaluated the validity of accelerometry as a passive digital phenotyping measure of negative symptoms in outpatients with SZ. Results generally supported the validity of accelerometry, as indicated by: (1) Criterion Validity: lower mean and SD phone ACL scores in SZ than CN; (2) Criterion Validity: Greater mean and SD phone ACL scores during contexts where participants reported being in public and engaged in recreational activities; (3) Convergent Validity: significant inverse correlations between band ACL measures and clinically rated negative symptoms on the BNSS; (4) Convergent Validity: significant positive correlations between phone ACL measures and EMA self-reports of engagement in goal-directed activity; (5) Discriminant Validity: nonsignificant and low correlations between ACL phone and band measures and positive, disorganized, and total symptoms on the PANSS.

Collectively, these findings provide proof-of-concept that ACL is feasible for use as an objective measure of negative symptoms in research studies. We recommend the following when incorporating ACL into research studies on negative symptoms: (1) Collect both band and phone ACL data, as different patterns of group differences

and correlations emerged across these devices. Both methods of collection demonstrated some validity, and it is currently unclear why they diverged and had limited overlap. Our sample had difficulty pairing the band and phone, which resulted in some loss of band data; we suspect this was device specific and believe alternate bands may prove more feasible for participants to troubleshoot and implement consistently (see Raugh et al.⁴³) for a discussion of adherence of passive digital phenotyping); (2) Band and phone ACL data both require intensive data processing; to be implemented in clinical trials, pharmaceutical companies may want to partner with experts in academia. Care should be taken with regard to cleaning and data reduction for phone versus band data, as these require different approaches due to the effects of gravitational forces and other collection characteristics; (3) Devices used to collect ACL, apps, and metrics used for scoring should undergo extensive validation processes before being implemented in clinical trials as outcome measures. The 3V approach proposed by Goldsack et al.⁴⁴ represents one model for doing so; (4) Evaluate a range of ACL metrics that reflect both vigor of movement and variability of movement; (5) The combination of passive ACL data and EMA-based surveys (e.g., context reports related to goal-directed activity) may yield more comprehensive estimates of negative symptoms than either active or passive digital phenotyping methods alone. Similarly, the EMA survey data may allow researchers to hone in on specific real-world contexts where passive digital phenotyping might be most relevant (e.g., when participants report being in public or engaged in goal-directed activities).

Certain limitations should be considered. First, this study was designed as proof-of-concept. More extensive investigations are needed with larger samples. This is particularly true for band-based ACL, as a portion of our sample had missing data due to bluetooth pairing problems. It is unclear if the participants who were not able to produce band data are comparable to those who were able to troubleshoot the band and produce data for both devices. Future studies should continue to compare band and phone data. We suspect the observed discrepancies may occur for a variety of reasons (e.g., participants may not carry the phone on them continuously, differences in how the accelerometers register movement). Second, participants used smartphones and bands that were provided to them for the study. It is unclear whether different patterns of data might emerge when participants use their own devices, and additional work is needed to confirm that ACL data is comparable across different operating systems and phones. Third, participants received payment for completing surveys and a bonus for returning the smart phone and band; it is unclear whether payment affected data acquisition, adherence, or quality. Fourth, data was collected for only 6 days. It is possible that longer periods of data collection could produce different results. Future studies should examine this issue directly, as it holds relevance for clinical trials. Further, although participants were instructed to carry the study phone and wear the band, participants may have not always been able to adhere to these procedures; this time period would not adequately capture a participant's ACL. Finally, correlation coefficients were in the medium range, and lower than what might be expected to demonstrate strong convergent validity. Temporally proximal EMA measures tended to have better convergent validity than retrospective clinical rating scales with the passive ACL measures. This discrepancy likely reflects temporal resolution, precision of measurement, and/or methods variance rather than inadequate convergent validity.

Despite these limitations, findings suggest that markers of movement vigor and variability, such as those measured via accelerometry or actigraphy⁴⁵⁻⁴⁷, hold promise as objective passive digital phenotyping measures of negative symptoms in SZ. Using this same sample, we previously demonstrated that SZ participants report high tolerability when completing active and passive digital phenotyping. However, feasibility and adherence

differed based on device. In particular, adherence (the % data received that was expected) was much lower for ACL measured via band (20%) than phone (87%) in SZ⁴³. Similarly, feasibility of the band was also poorer than the phone, with SZ participants reporting difficulties pairing the band and phone via bluetooth, whereas phone ACL demonstrated minimal feasibility issues (e.g., remembering to carry the phone). When coupled with the validity findings reported above, our prior tolerability/adherence/feasibility findings⁴³ suggest that while ACL may hold promise, methodological considerations are warranted. Until further validity is demonstrated and it is determined whether band or phone collection (or both in tandem) is most ideal, clinical trials may want to use ACL as a secondary or exploratory outcome measure. We believe it could be particularly useful when added to multi-level models that incorporate passive digital phenotyping, active digital phenotyping, and clinical symptom interviews at different levels, allowing for a more comprehensive evaluation of sensitivity to change across time.

METHODS

Participants

Data was collected from 50 individuals with DSM-5 diagnoses of schizophrenia or schizoaffective disorder (SZ) and 70 psychiatrically and neurologically healthy control (CN) participants. These groups did not differ on age, sex, parental education, or race; however, SZ had lower personal education and completed fewer EMA surveys than CN (see Table 3). Participants were not excluded on the basis of EMA survey completion adherence to maximize generalizability of ACL findings and because EMA surveys were not the core focus of the study. A full report of ACL and EMA adherence, feasibility, and tolerability is provided in Raugh et al.⁴³. However, we have not previously reported on the validity of ACL as an objective measure of negative symptoms.

SZ were recruited from local community mental health centers and online or printed advertisements. Diagnosis was made using the Structured Clinical Interview for DSM-5 (SCID)⁴⁸. CN were recruited from the local community using printed and online advertisements. CN had no current major psychiatric diagnoses as determined via the SCID-5 (i.e., mood, anxiety, substance, eating, compulsive, traumatic, somatic), no current schizophrenia-spectrum personality disorders as determined via the SCID-PD⁴⁹, no lifetime history of psychotic or bipolar disorders, no family history of psychosis, and were not currently prescribed any psychotropic medications. All participants reported being free from lifetime neurological disorders. Written informed consent was obtained from all participants for a protocol approved by the University of Georgia Institutional Review Board.

Variable	SZ (n = 50)	CN (n = 70)	Test Statistic
Age	38.7 (12.3)	34.7 (12.1)	$F = 3.1$
% Female	66%	71.4%	$\chi^2 = 0.40$
Personal education	13.2 (2.2)	15.1 (2.6)	$F = 17.2^{***}$
Parental education	13.7 (2.9)	14.2 (2.8)	$F = 0.86$
Race			$\chi^2 = 6.9$
Black	32%	21.4%	
Asian-American	0%	7.1%	
Biracial	6%	4.3%	
White	58%	57.1%	
LatinX	4%	7.1%	
Other	0%	2.9%	
Survey Adherence	57%	71%	$F = 10.3^{**}$

CN control, SZ schizophrenia.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Procedures

Phase 1: Initial laboratory visit. On the initial visit, participants provided informed consent and completed diagnostic/symptom interviews and digital phenotyping training procedures. Diagnostic/symptom interviews for SZ consisted of the SCID-5, Brief Negative Symptom Scale (BNSS)⁶, Positive and Negative Symptom Scale (PANSS)⁵⁰, and Level of Functioning Scale (LOF)⁵¹. Diagnostic interviews for CN consisted of SCID-5⁴⁸ and SCID-PD⁴⁹. Digital phenotyping training included instructions on how to use the study phone and band, complete surveys, and manage basic troubleshooting. Participants were also instructed to keep the phone on their person. To ensure participants understood the EMA and ACL procedures, participants completed a practice EMA survey and wore the band while receiving instructions on its use and how to troubleshoot. Additionally, participants were sent home with written instructions to help with key questions or technical problems (including Bluetooth connectivity) and told to call the researchers if there were any problems. All participants were called on the first day of Phase 2 to ensure proper survey delivery and troubleshoot any issues with the band.

All participants were provided with a Blu Vivo 5R Android smartphone programmed with all surveys for Phase 2 and the Embrace band (<https://www.empatica.com/embrace/>). Phone functions not relevant to study procedures were blocked. Of note, consistent with prior studies^{40,52,53}, we provided participants with a phone to ensure comparability of ACL data (e.g., ACL data may be calculated differently across phone operating systems or types).

Phase II: Digital phenotyping data collection. Surveys were preprogrammed and delivered using the mEMA application from Ilumivu (<https://ilumivu.com/>). In line with prior EMA studies⁵⁴⁻⁵⁶, surveys were scheduled to occur at eight quasi-random times between 9:00 and 21:00. A survey could occur at any time within a 90 min epoch with a minimum of 18 min and a maximum of three hours between surveys. Participants were notified of surveys by a tone and vibration (both of which could be turned off) at the scheduled time then five and ten minutes later if the survey was not completed. After 15 min, surveys became unavailable.

Surveys collected contextual variables, including current location ("Where are you?"), activity ("What are you doing?"), and social interaction ("Who are you interacting with?") (see Supplemental Materials for response options). As participants were instructed to include activities from the 15 min preceding the survey, multiple locations, activities, and interactions could be selected. A variable delineating performance of goal-directed activities was calculated for each survey instance³⁶. Specifically, we a priori identified the following survey responses as motivated or goal-directed activities: working/studying, errands/housework, exercise, shopping, or commuting/traveling. We then calculated the percentage of their EMA surveys where they were engaged in one of these goal-directed activities to achieve an index of the proportion of time participants spent in goal-directed activities. Of note, this measure of goal-directed activity encompasses but is more broad than prior papers that have used EMA to assess physical activity⁵⁷ but largely aligns with prior EMA studies' categorizations of effortful behavior⁵⁸ and productive activities⁵⁹ in severe mental illness samples.

Phone sensors were programmed to collect accelerometry values with each change in XYZ coordinate motion (every change in accelerometry being logged as a single instance), with separate values output for X, Y, and Z movement axes. The Embrace band collected accelerometry as gravitational force (g units) at a rate of 32 Hz in a range between -16g and 16 g. Embrace data was transferred to the phone via Bluetooth connection. If the connection was not available, the Embrace band could store up to 14 h of data. All data was encrypted and stored using unique identification codes on the Ilumivu or Empatica servers, separate from identifying information, until downloaded by the research team.

After 6 days of digital phenotyping data collection, participants returned the Embrace band, smartphone, and chargers and were compensated with \$20 per hour for completing study interviews. Following prior methods^{55,60}, participants were also compensated \$1 for each EMA survey completed. Participants additionally received an \$80 bonus for returning all study equipment.

Data preparation and analysis

Accelerometry (ACL) collected by the Embrace band was aggregated into minute epochs. As recording was in g's and XYZ angular coordinates were not recorded, continuous force of gravity is a confound in band data and cannot be removed. As such, g was converted to meters per second

squared ($1\text{ g} = 9.8\text{ m/s}^2$) to ensure consistency in units. Due to glitches with the smartphones, some instances had a recorded acceleration over 30 m/s^2 ($\sim 67\text{ mph}^2$, three times the force of gravity and outside of standard human accelerometry); all such instances were excluded⁶¹. Axial acceleration for the phone and band were summarized as the root of the sum of the squared values.

Phone and band accelerometry data was cleaned to remove outlier values. Outliers were identified separately for each group and defined as accelerometry values greater than two times the interquartile range. Less than .1% of band accelerometry samples and 8% of phone accelerometry samples were identified as outliers and removed.

Data was analyzed at the aggregate summary level across the entire week to fall on the same timeframe as clinical rating scales. Aggregate means (ACL mean) and standard deviations (ACL SD) were calculated for each subject for phone and band data. A measure of mean "activity index" was also calculated from the band data based on an algorithm from Bai et al.⁶². This measure reflects the amount of stationary observations.

To evaluate hypothesis 1, group differences in each phone and band accelerometry measure were compared using one-way ANOVA. Bivariate correlations were used to evaluate convergent and discriminant validity for hypotheses 2 and 3. Two-tailed tests were used.

DATA AVAILABILITY

All data is publicly available within the National Institute of Mental Health Data Archive (NDA) and can be accessed upon request.

CODE AVAILABILITY

All relevant analysis scripts are available upon request.

Received: 6 July 2021; Accepted: 24 February 2022;

Published online: 15 April 2022

REFERENCES

- Fervaha, G., Foussias, G., Agid, O. & Remington, G. Impact of primary negative symptoms on functional outcomes in schizophrenia. *Eur. Psychiatry* **29**, 449–455 (2014).
- Strauss, G. P., Harrow, M., Grossman, L. S. & Rosen, C. Periods of recovery in deficit syndrome schizophrenia: a 20-year multi-follow-up longitudinal study. *Schizophr. Bull.* **36**, 788–799 (2010).
- Strauss, G. P., Sandt, A. R., Catalano, L. T. & Allen, D. N. Negative symptoms and depression predict lower psychological well-being in individuals with schizophrenia. *Compr. Psychiatry* **53**, 1137–1144 (2012).
- Fusar-Poli, P. et al. Treatments of negative symptoms in schizophrenia: meta-analysis of 168 randomized placebo-controlled trials. *Schizophr. Bull.* **41**, 892–899 (2015).
- Kirkpatrick, B., Fenton, W. S., Carpenter, W. T. & Marder, S. R. The NIMH-MATRICES consensus statement on negative symptoms. *Schizophr. Bull.* **32**, 214–219 (2006).
- Kirkpatrick, B. et al. The brief negative symptom scale: psychometric properties. *Schizophr. Bull.* **37**, 300–305 (2011).
- Kring, A. M., Gur, R. E., Blanchard, J. J., Horan, W. P. & Reise, S. P. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. *Am. J. Psychiatry* **170**, 165–172 (2013).
- Strauss, G. P. et al. Next-generation negative symptom assessment for clinical trials: validation of the brief negative symptom scale. *Schizophr. Res.* **142**, 88–92 (2012).
- Strauss, G. P. et al. Factor structure of the brief negative symptom scale. *Schizophr. Res.* **142**, 96–98 (2012).
- Strauss, G. P., Vertinski, M., Vogel, S. J., Ringdahl, E. N. & Allen, D. N. Negative symptoms in bipolar disorder and schizophrenia: a psychometric evaluation of the brief negative symptom scale across diagnostic categories. *Schizophr. Res.* **170**, 285–289 (2015).
- Strauss, G. P. & Gold, J. M. A psychometric comparison of the clinical assessment interview for negative symptoms and the brief negative symptom scale. *Schizophr. Bull.* **42**, 1384–1394 (2016).
- Strauss, G. P. & Chapman, H. C. Preliminary psychometric properties of the brief negative symptom scale in youth at clinical high-risk for psychosis. *Schizophr. Res.* **193**, 435–437 (2018).
- Ahmed, A. O. et al. Cross-cultural validation of the 5-factor structure of negative symptoms in schizophrenia. *Schizophr. Bull.* **45**, 305–314 (2019).
- Forbes, C. et al. Initial development and preliminary validation of a new negative symptom measure: the clinical assessment interview for negative symptoms (CAINS). *Schizophr. Res.* **124**, 36–42 (2010).
- Horan, W. P., Kring, A. M., Gur, R. E., Reise, S. P. & Blanchard, J. J. Development and psychometric validation of the clinical assessment interview for negative symptoms (CAINS). *Schizophr. Res.* **132**, 140–145 (2011).
- Bischof, M. et al. The brief negative symptom scale: validation of the German translation and convergent validity with self-rated anhedonia and observer-rated apathy. *BMC Psychiatry* **16**, 415 (2016).
- Chan, R. C. K. et al. Validation of the Chinese version of the clinical assessment interview for negative symptoms (CAINS): a preliminary report. *Front. Psychol.* **6**, 7 (2015).
- Chieffi, M. et al. The brief negative symptom scale: convergent/discriminant validity and factor structure in a large sample of outpatients with schizophrenia. *Eur. Psychiatry* **30**, 246 (2015).
- de Medeiros, H. L. V. et al. The brief negative symptom scale: validation in a multicenter Brazilian study. *Compr. Psychiatry* **85**, 42–47 (2018).
- Engel, M., Fritzsche, A. & Lincoln, T. M. Validation of the German version of the clinical assessment interview for negative symptoms (CAINS). *Psychiatry Res.* **220**, 659–663 (2014).
- García-Portilla, M. P. et al. Psychometric evaluation of the negative syndrome of schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* **265**, 559 (2015).
- Hosáková, K., Viktorová, L. & Lečbych, M. New methods for diagnostics of negative symptoms in schizophrenia: clinical assessment interview for negative symptoms (CAINS)-validation of psychometric properties of the Czech version. *Cesk Psychol.* **61**, 245–256 (2017).
- Jang, S.-K. et al. Validation of the Korean version of the clinical assessment interview for negative symptoms. *Psychiatry Investig.* **14**, 413–419 (2017).
- Jung, S. I., Woo, J., Kim, Y.-T. & Kwak, S. G. Validation of the Korean-version of the clinical assessment interview for negative symptoms of schizophrenia (CAINS). *J. Korean Med. Sci.* **31**, 1114–1120 (2016).
- Mané, A. et al. Spanish adaptation and validation of the brief negative symptoms scale. *Compr. Psychiatry* **55**, 1726–1729 (2014).
- Merlotti, E., Mucci, A., Bucci, P., Nardi, A. & Galderisi, S. Italian version of the "brief negative symptom scale". *J. Psychopathol. Behav. Assess.* **20**, 199–215 (2014).
- Mucci, A. et al. The brief negative symptom scale (BNSS): independent validation in a large sample of Italian patients with schizophrenia. *Eur. Psychiatry* **30**, 641–647 (2015).
- Polat Nazlı, I. et al. Validation of Turkish version of brief negative symptom scale. *Int. J. Psychiatry Clin. Pract.* **20**, 265–271 (2016).
- Valiente-Gómez, A. et al. Validation of the Spanish version of the clinical assessment for negative symptoms (CAINS). *Schizophr. Res.* **166**, 104–109 (2015).
- Wójcicki, P. et al. Polish version of the brief negative symptom scale (BNSS). *Psychiatr Pol.* **53**, 541–549 (2019).
- Xie, D. et al. Cross cultural validation and extension of the clinical assessment interview for negative symptoms (CAINS) in the Chinese context: evidence from a spectrum perspective. *Schizophr. Bull.* **44**, S547–S555 (2018).
- Yao et al. Reliability and validity of the Chinese version of brief negative symptom scale. *Chin. Ment Health J.* **28**, 302–307 (2014).
- Strauss, G. P. & Gold, J. M. A new perspective on anhedonia in schizophrenia. *Am. J. Psychiatry* **169**, 364–373 (2012).
- Insel, T. R. Digital phenotyping: technology for a new science of behavior. *JAMA* **318**, 1215–1216 (2017).
- Onnela, J.-P. & Rauch, S. L. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* **41**, 1691–1696 (2016).
- Raugh, I. M. et al. Geolocation as a digital phenotyping measure of negative symptoms and functional outcome. *Schizophr. Bull.* **46**, 1596–1607 (2020).
- Yates, T. et al. Impact of depression and anxiety on change to physical activity following a pragmatic diabetes prevention program within primary care: pooled analysis from two randomized controlled trials. *Diabetes Care* **42**, 1847–1853 (2019).
- Firth, J. et al. The validity and value of self-reported physical activity and accelerometry in people with schizophrenia: a population-scale study of the UK Biobank. *Schizophr. Bull.* **44**, 1293–1300 (2018).
- Ben-Zeev, D., Scherer, E. A., Wang, R., Xie, H. & Campbell, A. T. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. *Psychiatr Rehabil. J.* **38**, 218–226 (2015).
- Ben-Zeev, D. et al. CrossCheck: integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatr Rehabil. J.* **40**, 266–275 (2017).
- Wang, R. et al. Predicting symptom trajectories of schizophrenia using mobile sensing. *IMWUT* **1**, 1–24 (2017).

42. Blanchard, J. J., Kring, A. M., Horan, W. P. & Gur, R. Toward the next generation of negative symptom assessments: the collaboration to advance negative symptom assessment in schizophrenia. *Schizophr. Bull.* **37**, 291–299 (2011).
43. Raugh, I. M. et al. Digital phenotyping adherence, feasibility, and tolerability in outpatients with schizophrenia. *J. Psychiatr. Res.* **138**, 436–443 (2021).
44. Goldsack, J. C. et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *npj digital. npj Digital Med.* **3**, 1–5 (2020). Apr 14.
45. Umbricht, D., Cheng, W.-Y., Lipsmeier, F., Bamdadian, A. & Lindemann, M. Deep learning-based human activity recognition for continuous activity and gesture monitoring for schizophrenia patients with negative symptoms. *Front. Psychiatry* **11**, e574375 (2020).
46. Krane-Gartiser, K., Henriksen, T. E., Morken, G., Vaaler, A. E. & Fasmer, O. B. Motor activity patterns in acute schizophrenia and other psychotic disorders can be differentiated from bipolar mania and unipolar depression. *Psychiatry Res.* **270**, 418–425 (2018).
47. Wee, Z. Y. et al. Actigraphy studies and clinical and biobehavioural correlates in schizophrenia: a systematic review. *J. Neural Transm.* **126**, 531–558 (2019).
48. First M. B. *Structured clinical interview for the DSM (SCID). The encyclopedia of clinical psychology*, (John Wiley & Sons, Inc, 2014), p. 1–6. <https://doi.org/10.1002/9781118625392.wbecp351>
49. First M. B., Williams J. B., Benjamin L. S., Spitzer R. L. *SCID-5-PD: structured clinical interview for DSM-5® personality disorders*. (American Psychiatric Association Publishing; 2016).
50. Kay, S. R., Fiszbein, A. & Opler, L. A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
51. Hawk, A. B., Carpenter, W. T. & Strauss, J. S. Diagnostic criteria and five-year outcome in schizophrenia: a report from the international pilot study of schizophrenia. *Arch. Gen. Psychiatry* **32**, 343–347 (1975).
52. Ben-Zeev, D. et al. Mobile behavioral sensing for outpatients and inpatients with schizophrenia. *Psychiatr. Serv.* **67**, 558–561 (2016).
53. Fulford, D. et al. Smartphone sensing of social interactions in people with and without schizophrenia. *J. Psychiatr. Res.* **137**, 613–620 (2021).
54. Depp, C. A. et al. Social behavior, interaction appraisals, and suicidal ideation in schizophrenia: the dangers of being alone. *Schizophrenia Res.* **172**, 195–200 (2016).
55. Granholm, E. et al. What do people with schizophrenia do all day? ecological momentary assessment of real-world functioning in schizophrenia. *Schizophr. Bull.* **46**, 242–251 (2020).
56. Mote, J. & Fulford, D. Ecological momentary assessment of everyday social experiences of people with schizophrenia: a systematic review. *Schizophr. Res.* **216**, 56–68 (2020).
57. Andorko, N. D. et al. Assessing validity of retrospective recall of physical activity in individuals with psychosis-like experiences. *Psychiatry Res.* **273**, 211–217 (2019).
58. Gard, D. E. et al. Do people with schizophrenia have difficulty anticipating pleasure, engaging in effortful behavior, or both? *J. Abnorm. Psychol.* **123**, 771–782 (2014).
59. Strassnig, M. T. et al. Evidence for avolition in bipolar disorder? A 30-day ecological momentary assessment comparison of daily activities in bipolar disorder and schizophrenia. *Psychiatry Res.* **300**, 113924 (2021).
60. Strassnig, M. T., Harvey, P. D., Miller, M. L., Depp, C. A. & Granholm, E. Real world sedentary behavior and activity levels in patients with schizophrenia and controls: an ecological momentary assessment study. *Ment. Health Phys. Act.* **20**, 100364 (2021).
61. Chen, K. Y. & David R Bassett, J. R. The technology of accelerometry-based activity monitors: current and future. *Med. Sci. Sports Exerc.* **37**, S490–S500 (2005).
62. Bai, J. et al. An activity index for raw accelerometry data and its comparison with other activity metrics. *Plos One* **11**, e0160644 (2016).

ACKNOWLEDGEMENTS

The authors wish to thank the participants who completed the study. Research was supported by USA National Institute of Mental Health grant R21-MH112925 to Dr. Strauss. Study sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

COMPETING INTERESTS

G.P.S. and B.K. are original developers of the Brief Negative Symptom Scale (BNSS) and receive royalties and consultation fees from Medavante-ProPhase LLC in connection with commercial use of the BNSS and other professional activities; these fees are donated to the Brain and Behavior Research Foundation. B.K. has received honoraria and travel support from ProPhase LLC for training pharmaceutical company raters on the BNSS, consulting fees and travel support from Genentech/Roche, Minerva Neurosciences, Lundbeck, and ProPhase LLC, consulting fees from anonymized pharmaceutical companies through Decision Resources, Inc, and from an investment capital company through Guideposts, and from Wockhardt Bio AG for consulting on a legal issue. G.P.S. has consulted for Minerva Neurosciences, Acadia, Otsuka, Boehringer-Ingelheim, and Lundbeck. GPS, BK, and ASC are cofounders of Quantic Innovations LLC that provides resources for digital phenotyping data collection, analysis, and interpretation. All other authors have no conflicts to report.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41537-022-00241-z>.

Correspondence and requests for materials should be addressed to Gregory P. Strauss.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022