10-2022

Relevance of digital measures to people with early Parkinson's for monitoring meaningful symptoms in the WATCH-PD study

Jennifer R. Mammen
University of Rhode Island, jmammen@uri.edu

et al.

Citation/Publisher Attribution

Mammen, J.R., Speck, R., Stebbins, G., Müller, M., Yang, P., Campbell, M., . . . Adams, J.L. (2022). Relevance of digital health technologies to people with early Parkinson's for monitoring meaningful symptoms in the WATCH-PD study. (Preprint).https://digitalcommons.uri.edu/nursing_facpubs/345/

Related manuscripts in this collection:

1. (METHODS) Mammen, J.R., Speck, R., Stebbins, G., Müller, M., Yang, P., Campbell, M., . . . Adams, J.L. (2022). Mapping meaningful symptoms and impacts of disease to digital outcome measures. (Preprint). Retrieved from https://digitalcommons.uri.edu/nursing_facpubs/343/
2. (RESULTS) Mammen, J.R., Speck, R., Stebbins, G., Müller, M., Yang, P., Campbell, M., . . . Adams, J.L. (2022). Relative meaningfulness and impacts of symptoms in people with early-stage Parkinson's disease. (Preprint). Retrieved from https://digitalcommons.uri.edu/nursing_facpubs/344/
3. (INSTRUMENT). Mammen, J.R., & Adams, J.L. (2022). Clinical Outcome Assessment Meaningfulness Survey (COA-MeanS).

Relevance of digital measures to people with early Parkinson's for monitoring meaningful symptoms in the WATCH-PD study

Authors
Jennifer R. Mammen, PhD[1]
Rebecca M. Speck, PhD, MPH[2]
Glenn T. Stebbins, PhD[3]
Martijn L.T.M. Müller, PhD[2]
Phillip T. Yang[4]
Michelle Campbell, PhD[9]
Josh Cosman, PhD[6]
John E. Crawford[a]
Tien Dam, MD[12]
Johan Hellsten, PhD[7,a]
Stella Jensen-Roberts, MD[4]
Melissa Kostrzebski, MBA[4,5]
Tanya Simuni, MD[8]
Kimberly Ward Barowicz[2]
Jesse M. Cedarbaum, MD[10, 11]
E. Ray Dorsey, MD[4,5]
Diane Stephenson, PhD[2]
Jamie L. Adams, MD[4,5]

[a]person with Parkinson's (PwP)

Affiliations:
1. University of Rhode Island, College of Nursing
2. Critical Path Institute, Tucson, AZ, USA
3. Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA
4. Center for Health + Technology, University of Rochester Medical Center, Rochester, NY, USA
5. Department of Neurology, University of Rochester, Medical Center, Rochester, NY, USA
6. Abbvie Inc., North Chicago, USA
7. H. Lundbeck A/S, Valby, Denmark
8. Northwestern University Feinberg School of Medicine, Chicago IL
9. Center for Drug Evaluation and Research (CDER), U.S. Food and Drug Administration (FDA)
10. Coeruleus Clinical Sciences LLC, Woodbridge CT
11. Yale Medical School, New Haven CT
12. Biogen, Cambridge USA

Corresponding Author: Jennifer Mammen
jmammen@uri.edu
University of Rhode Island College of Nursing,
350 Eddy Street, Providence RI

Word Counts

Title:
Number of references: 31
Number of tables: 4
Number of figures: 3
Supplemental online materials: 3
Word count abstract: 250
Word count paper: 4000

Keywords: Parkinson's, digital health technology, qualitative, meaningfulness

Funding Source: This study was funded by the Critical Path for Parkinson's (CPP) Consortium. The CPP 3DT initiative is funded by the CPP Consortium members including: Biogen; GSK; Takeda; Lundbeck; UCB Pharma; Roche; AbbVie and Merck, Parkinson's UK, and the Michael J Fox Foundation. Critical Path Institute is supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) and is 54.2% funded by the FDA/HHS, totaling $13,239,950, and 45.8% funded by non-government source(s), totaling $11,196,634. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS or the U.S. Government.

Quality in qualitative research statement

This study used COREQ criteria to guide reporting of qualitative findings.

Table legends

Table 1. WATCH-PD Digital measures evaluated for relevance to personally meaningful symptoms from the participant perspective

Table 2. Demographic characteristics of the sample comparative to the parent WATCH-PD study

Table 3. Frequency and percentage of positive demonstration of understanding and relevance of WATCH-PD tasks during cognitive interviewing vs. mapping

Table 4. Theme 1 supporting data - Examples of task endorsement based on five criteria used by participants to explain rationales for personal relevance

Figure legends

Figure 1. Sample symptom map with pictographs connecting WATCH-PD digital measure to personally important symptoms

Figure 2. Perceived relevance of WATCH-PD Tasks to monitoring any personally important (experienced or anticipated) versus bothersome symptoms in people with early PD as elicited by Mapping

Figure 3. Recommended approach to assessing the relevance of digital measures for monitoring meaningful symptoms of disease

Supplemental Figure and Table Legends (Online only)

Supplement A. Coding Schema for assessing relevance of digital measures to meaningful symptoms

Supplement B. Supporting data for Theme 2

Supplement C. Supporting data for Theme 3

Abstract

Background: Adoption of new digital measures for clinical trials and practice has been hindered by lack of data demonstrating relevance of these metrics to people with Parkinson's.

Objective: This study aimed to explore participant-perceived relevance of WATCH-PD digital measures to monitoring personally important aspects of disease in people with early PD.

Methods: Forty participants from the WATCH-PD study were solicited for 1:1 in-depth online interviews. Cognitive interviewing was combined with symptom mapping techniques to assess content validity of 10 digital measures and relevance to personal symptoms within a rank-ordered symptom hierarchy. Content analysis and descriptive techniques were used to analyze data.

Results: Most tasks (9/10) were rated relevant to personally important symptoms of PD by both cognitive interviewing (70—92.5%) and mapping (80—100%). However, only two related to currently bothersome symptoms for more than 80% of participants (*Tremor*, *Shape Rotation)*. Tasks were generally deemed relevant if they met three participant context criteria: (1) understanding what the task measured, (2) believing it targeted an important symptom of PD (past, present, or future), and (3) believing the task was a good test of that important symptom. Participants did not require that a task relate to active symptoms or "real" life to be relevant.

Conclusion: In this study of individuals with early PD, participants rated digital measures of tremor and hand dexterity as most relevant to them. Thematic findings further showed that participants perceptions of relevance of tasks were based on person-centric criteria not typically included in standard content validity assessments.

Background

Digital measures appear to offer unparalleled opportunities to detect early changes in symptoms and disease progression, which could support the development of urgently needed disease modifying therapies for neurodegenerative disorders such as Parkinson's disease (PD).[1,2] However, adoption of such measures for clinical outcomes assessments in therapeutic trials has been hindered by limited patient experience data, including inability to transparently demonstrate relevance of these technologies to patients' experience of living with the disease.[3-5]

Rapidly growing regulatory emphasis by the U.S. Food and Drug Administration on patient-focused drug development has translated to an urgent need to efficiently evaluate the extent to which new technologies align with meaningful aspects of disease from the patient perspective.[6,7] However, this objective has proved challenging under traditional evaluation approaches, with quantitative techniques (i.e surveys, rating scales) being generally inadequate to reflect the depth of personal experiences, and qualitative techniques (i.e. interviews) often lacking the consistency and precision needed to quantify prevalence of opinions and experiences.[8] Thus, we elected a mixed-methods approach combining new symptom mapping approaches [9] with standardized cognitive interviewing techniques[10] to assess the relevance of using smartwatch and smartphone applications[11] for monitoring meaningful aspects of Parkinson's disease from the perspective of people with early PD.

Methods

Study background

This qualitative study was conducted in follow-up to the WATCH-PD study (NCT03681015), which was a 12-month multi-center observational trial that evaluated ability of novel smartwatch and smartphone applications to monitor symptoms and disease progression in people with early, untreated PD (≤ 2 years diagnosis, Hoehn & Yahr stage ≤ 2).[12] The current study focused on evaluating participant-perceived relevance of the WATCH-PD smartwatch and smartphone measures, which consisted of 10 activities (descriptions and images show in **Table 1)** that were completed during in-clinic visits and bi-weekly at home.

Sample

Individuals with PD who had completed their final visit of the WATCH-PD study within the previous 6 months were eligible for the current study (N=54).[13] All non-white participants (N=4) were solicited, with purposeful sampling that otherwise mirrored parent study demographics. IRB approval and digital informed consent was obtained (IRB#00003002).

Data collection

A multi-faceted data collection approach was used. This consisted of an online survey followed by a 1:1 virtual interview a week later, which included systematically mapping all important symptoms and impacts of PD, followed by formal cognitive interviewing on the 10 digital measures, and lastly mapping of the digital measures back to the personal symptoms in the symptom map. Surveys were used to gather information on demographics, PD medication use, self-reported PD symptoms, and preliminary qualitative data on perceived relevance of

WATCH-PD tasks to monitoring symptoms of PD. Online interviews were conducted by JM (white, female, PhD prepared nurse practitioner, unacquainted with participants) via Zoom video-conferencing and were audio/visually recorded with participant permission.

Part 1: Symptom mapping.[9] In the first part of the online interview, each participant was assisted to create a detailed map of their PD symptom experience. Symptom maps were created using Xmind™ mapping software, via Zoom screen sharing, with symptoms (represented as nodes) and descriptions of symptoms (represented as lines), with symptoms arranged hierarchically top to bottom and grouped by order of bothersomeness into five categories: "Most bothersome," "Somewhat bothersome," "Less bothersome," and "No current issues (but still important),". An additional bottom tier category was included for items viewed by the participant as "Not relevant to early PD".

***Part 2: Cognitive interviewing*.** After completing the symptom map, the interviewer conducted a structured assessment of understanding and perceived relevance of WATCH-PD tasks to monitoring personally important symptoms of PD. Four standardized questions were used to evaluate content validity: (Q1) Did the participant understand what the task was asking them to do in order to complete it? (Q2) Did completion of the task relate to the participant's PD symptoms? (Q3) Was the task similar or relevant to tasks or activities in the participant's daily life? (Q4) Did the participant perceive the task was relevant to measuring the progression of their PD?

Part 3: Mapping of task to personally important symptoms. Following cognitive interviewing, participants returned to their symptom map and were asked to incorporate small pictographs of each of the 10 tasks into their personal map building on the preceding discussion

of perceived relevance of that task to monitoring personally important symptoms. For this step, the participant placed the task image next to the personal symptom(s) they believed the task measured, or into the "Not relevant" category, as applicable. They were then asked to further reflect on and summarize what made each task relevant or not relevant, adding clarifying information as needed. A sample symptom map showing conjunction of tasks to personal symptoms is presented in **Figure 1.**

Data analysis

Descriptive statistics were computed for survey items. Coding of the four standardized cognitive interview questions assessing content validity was performed by RS, with frequencies and percentages calculated for each. Content coding[15] of symptom maps was performed by JM and PY (>97% convergence). Relevance of each WATCH-PD task was calculated based on direct association of the task with any *personally important symptom* in the symptom map. Important symptoms comprised all PD symptoms that were meaningful to the individual—past, present, or anticipated—with each experienced symptom further classified as more or less bothersome using Patient Reported Symptom Scores (PRSS; 4=Most bothersome, 3=Somewhat bothersome; 2=Less bothersome, 1=Present but not bothersome, 0=No issues but still important, and "."=not relevant to early PD), as shown in **Supplement A**. For example, if a participant placed the *Tremor Task* next to the symptom "Tremor" at the "most bothersome" position in the symptom map, it would be associated with a PRSS of 4. Conversely, if the *Tremor Task* was placed next to a symptom under the "No issues" category, it would have a PRSS of 0 (not present but still personally important). If associated with more than one symptom (e.g.,

Walking & Balance), tasks were valued at the highest level they occurred at within a map and counted once per participant. Independent T-tests were run using SPSS28 to assess for differences in relevance of tasks based on any/no PD medication use. Cramer's V (effect size for Chi-square test of independence) was calculated to assess strength of relationship between cognitive interviewing and mapping assessment of task relevance, with large effect (i.e. very strong relationship) defined as >0.25.[16] Thematic analysis[17,18] of verbatim transcripts was conducted using Nvivo12.[19]

Rigor & Validity. Data collection protocols and procedures were pretested with people with PD (JC, JH), and initial interviews were observed for consistency (RS). Data collection and analysis (including content validity assessment) was performed by researchers unaffiliated with the parent study to minimize bias. Other measures to enhance validity included systematic member-checking to ensure accuracy of data collected, regular peer-debriefing to discuss and refine emerging themes, use of multiple coders to increase validity of analysis, and a formal audit trail.[20] Quotes throughout are presented with numeric identifiers to demonstrate representativeness.

Data availability statement

Data are available to Critical Path for Parkinson's (CPP) Consortium 3DT Initiative Stage 2 members. Non-members may submit proposals for de-identified datasets to CPP 3DT via the corresponding author.

Results

Sample characteristics

Forty people completed surveys and interviews. Participants were predominantly male (52.5%), white (92.5%), college educated (100%), with no PD medication use (60%), a mean age of 64 years (SD=8.8) and 2.1 years since PD diagnosis (SD=0.9). **Table 2** displays demographics comparative to the parent study. Of the 14/54 eligible people who did not participate, 1 declined, 5 were unreachable, and 8 were not solicited due to having reached goal sample size of 40, reflective of parent study demographics.

Cognitive interviewing: Content validity

The frequencies and percentages of positive demonstration of understanding and relevance of WATCH-PD tasks, as assessed through four standardized questions are shown in **Table 3**. With the exception of *Finger Tapping* (1/40 incorrect), all participants were able to demonstrate that they understood what each task was asking them to do for completion. The task most relevant to personal PD symptoms by cognitive interviewing was *Shape Rotation* (87.5%), followed by *Tremor* (85%) and *Finger Tapping* (85%). Less than half of participants found *Visual Spatial Memory* (47.5%), *Phonation* (42.5%), *Reading* (37.5%), *Articulation* (37.5%), and *Verbal Symbol Swap* tasks (35%) to be relevant to personal PD symptoms. The majority of participants were able to provide examples of how each task related to activities in their daily life (range 59% [*Visual Spatial Memory*] to 97.5% [*Walking & Balance*]) and believed tasks were important for measuring the progression of PD (52.5—97.5%). Excluding *Visual Spatial Memory*

and *Phonation,* all other tasks (8/10) were definitively endorsed as personally relevant by greater than 85% of participants.

Symptom maps: Participant rated relevance of tasks to *bothersome* symptoms

As shown in **Figure 2** (**Graph A**), tasks that were relevant to actively bothersome symptoms for the largest number of participants were the *Tremor* task (92.5%; relevant to tremor), *Shape Rotation*, *Finger Tapping*, (82.5%, 77.5% respectively; relevant to fine motor) and *Trails A&B* (65%; perceived as relevant to fine motor and thinking), *Walking & Balance* (55%, relevant to general mobility), and *Verbal Symbol Swap* (52.5%, relevant to cognitive function). Speaking tasks (*Reading, Phonation, Articulation*) and *Visual Spatial Memory* were less often related to actively bothersome symptoms (<50%).

Symptom maps: Participant rated relevance of tasks to personally *important* symptoms

Personally important symptoms included symptoms that were *past* (resolved) *present* (more or less bothersome) or *anticipated* (i.e., symptoms participants were concerned could occur down the road). For example, a participant who did not currently have difficulty with articulation might perceive tasks that monitor speech as relevant to an important future symptom, as seen here:

> ***P6****: Well, I'm not experiencing the slurring of words, but that [digital measure] could've picked up on something like that. I want to speak clearly...I don't want to be slurring words, [and] I completely agreed with [testing] this… I absolutely believe it's valid.*

***P19**: I don't have an issue [with speech]. [But] I think the tasks are good to have in the test, because a lot of people with Parkinson's do have these issues. How would I know if I did or didn't have them without doing the task?*

As shown in **Figure 2 (Graph B)**, all tasks, excluding *Visual Spatial Memory*, were rated as relevant for monitoring personally important symptoms to the greater majority (>80%). While speaking tasks did not relate to actively bothersome symptoms for most, they were relevant to more than 80% of people as important future symptoms. *Visual Spatial Memory* was commonly criticized for being difficult to complete, causing frustration, and failing to measure symptoms specific to PD, and was therefore rated as relevant by fewer people (52.5%).

Correlations. As shown in **Table 3,** convergence between cognitive interviewing Q4 (relevance of task to monitoring progression of PD) and relevance to monitoring personally important symptoms rated by mapping was generally high (67.5 [*Visual spatial Memory*]—97.5% [*Tremor*]) with evidence of strong association between ratings of relevance across all tasks (Cramer's V >0.25). Non-convergence increased if participants were unsure what the test measured, or felt the test was technically flawed, with 50% lack of concordance accounted for by items where participants were unable to give a categorical response (yes/no) to Q4 (relevant to monitoring PD progression), as seen in these examples:

***P18**: "I don't know [if it's relevant], but "yes" if it's about volume" [Phonation].*

***P16**: "Yes" if it was designed properly [Visual Spatial Memory]*

In these instances, participants often modified ratings of relevance between methods, suggesting that variability was caused by ambiguity in perception of the task.

Themes

Three themes were identified relating to how participants perceived the relevance of WATCH-PD tasks for monitoring early PD. In general, participants felt that digital measures were meaningful and should be retained for monitoring *if* they reliably captured any important symptoms of PD—whether actively experienced or not (Themes 1 and 2). However, there was some concern about negative psychological impact related to frequent monitoring of PD symptoms (Theme 3).

Theme 1. *Perceived relevance of digital measures was contingent on belief that the measure effectively evaluated an important PD symptom—regardless of whether the symptom was currently present or the measure related to activities in daily life.*

Relevant measures fell into two categories: (1) those that appeared related to personal symptoms actively experienced and (2) those that appeared related to personally important symptoms of Parkinson's *not* currently experienced by the participant (i.e., past or anticipated possible future symptoms of concern or common symptoms of PD). **Table 4** presents a detailed delineation of the five criteria on which participants consistently based their evaluation of personal relevance of WATCH-PD measures during interviews (hereafter called "Participant Criteria"). In general, participants endorsed measures as relevant if three Participant Criteria were met—the individual believed (1) they understood what was being evaluted (*purpose of measure*), (2) the measure related to an *important symptom* of PD, and (3) the measure was a *good test* of the symptom. Often, WATCH-PD measures were considered relevant even when they did not relate to activities in daily life or were a personally experienced symptom, provided

other Participant Criteria were met. Understanding what the measure required to complete it was not a referenced factor in personal relevance evaluations, with participants most focused on the purpose of the measure and whether they thought it was a good/bad test of the targeted PD symptom (i.e., Participant Criteria 1-3).

Some slight uncertainty was tolerated regarding what the measure was targeting (Participant Criteria 1), but if measures failed to satisfy Participant Criteria 2 and 3 (related to important symptoms and a good test), they were always deemed not relevant. However, participants indicated ambiguous measures could *become relevant* if data demonstrated ability to pick up subtle variations in important PD symptoms not apparent otherwise, as seen in the following quotes:

> ***P3****: I don't know what they were measuring, so it's hard for me to know whether it was related to my vocal quality. If it is able to measure the things I care about, yes [it's important]. It's hard for me to know without seeing the data.*

> ***P5****: I don't feel like I have enough information [to decide if the measure is relevant]. If I was having trouble completing [the measure], then I would say it was due to Parkinson's symptoms…but it's not something I've noticed. It would be nice to see results saying, "Wow, his walking has changed." [The digital measures could be] seeing things that I'm not.*

Theme 2. *Due to the variable, progressive, and uncertain nature of PD, people with early PD believe it can be important to measure aspects of disease and symptoms that they do <u>not</u> currently experience.*

Many participants indicated that they felt it was important to measure common symptoms of PD even when they did not personally experience those symptoms (Criteria #5). Rationales for this position were based on participant's understanding that PD has wide variability in symptom expression and disease progression, and that new symptoms could onset at any point. Participants believed that digital measures had the potential to pick up subtle deficits and new symptoms at the earliest possible onset and felt that earliest detection was key to understanding and monitoring disease progression and developing better future treatments. Furthermore, many reported being personally committed to partnering with researchers to find better treatments for the sake of themselves, family members who might later develop PD, and the broader benefit of society as a rationale for monitoring symptoms that were not yet present. This is illustrated in the following quotes with supporting data in **Supplement B**:

> ***P24****: It's not so much what you have currently, it's the progression. It's about whether new symptoms develop and if those symptoms become more severe over time. …Symptoms change, they get worse, or they suddenly show up, and you didn't have it before.*
>
> ***P28****: It's a degenerative disease… I think it's important to measure all of these [symptoms].*

Theme 3. *Active monitoring can have negative emotional impact that affects sense of well-being and engagement with the measurement tool, particularly if challenging to perform.*

Participants reported that when they perceived themselves as not performing well, they sometimes experienced negative emotional consequences, such as anxiety, stress, worry about disease progression, or sense of failure. For example:

***P26**: I felt like oh no, I can't handle this. This is gonna make me feel so inadequate. ...After I was all done, I would go be depressed for an hour because it pointed out to me that I was having trouble with some of these things.*

This sense of failure resulted in some participants being less willing to complete a task, approaching the activity with anxiety, finding ways to "game" the system to improve scores and sense of competence, and practicing to improve performance. Supporting data for this is presented in **Supplement C**, and is illustrated in the following examples:

***P16**: It's designed to dispirit people taking it. It caused me anxiety. By the end, people are just so frustrated they'll tap any finger. They won't care anymore.*

***P20**: I tried to focus to the best of my ability, but it became frustrating, and I would just say, "oh, I'll just hit match or no match… I'll just guess at that."*

***P27**: It motivated me to download [a different] app on my phone to try to get better at color recognition and [prove] that maybe my brain is not totally declined.*

Conversely, when participants were easily able to perform a task, or perform at a level they felt was equivalent to a person without PD, they experienced a sense of relief and felt more positively about the digital measure and their PD state. For this reason, along with desire to detect earliest symptom onset, many people preferred to retain items that measured important symptoms of PD not currently present, as seen here:

***P28**: If it was all hard, I might have dropped out on the study, because you're making me face this thing that's really, really, really hard to face. [But if] I can do at least half the stuff well, and they're measuring something that has to do with Parkinson's... It's like it's a mental boost – it's reassuring. I need some pats on the back. Please don't [take the easy tests out].*

Discussion

This is the first study to use symptom mapping with cognitive interviewing to systematically explore relevance of using digital measures for monitoring meaningful symptoms from the perspectives of people with early PD. Overall, nine of the ten WATCH-PD measures were rated as relevant to monitoring personally important symptoms of PD by both cognitive interviewing (70—92.5%) and mapping methodologies (80—100%). However, only two related to actively bothersome symptoms for more than 80% of the sample (*Tremor*, *Shape Rotation)*.

Data from mapping correlated well with and validated cognitive interview findings but also extend current understandings of DHT relevance within two key contexts: *actively bothersome* versus *personally important* symptoms (past, present or future). Notably, our data suggest that eliminating measures based on lack of current symptoms may not align with patient preferences for monitoring, as was evident in the higher rates of reported relevance vs. experienced symptoms. For instance, nearly all speech tests were rated as relevant (>80%) despite few reporting symptoms in this area (<50%), with repeatedly expressed preference to proactively monitor for these high-impact symptoms. Thematic findings also highlighted preferences to monitor for actively present *and* important future symptoms (i.e., those not yet experienced). Thus, digital assessment of common impactful symptoms might be warranted prior to onset to support early detection, trending, and treatment, as has been suggested in previous works.[25,26] We would propose that the distinction between what is actively bothersome versus what is important represents a critical distinction in classification of relevance of digital measures from the participant perspective, and should be considered in future assessments.

A second important finding of the study was that "relevance" of measures from the participant perspective depended on person-centric criteria that extended beyond traditional cognitive interviewing approaches to content validity assessment. This finding aligns with increasing regulatory recognition of the need to prioritize patient perspectives in the selection of clinical outcome assessments.[6,21-23] In this early PD sample, we found that concepts of relevance were contingent upon three participant context criteria being met: (1) believing they understanding what was measured, (2) believing it targeted an important symptom of PD (past, present, or future), and (3) believing the measure was a good test of that important symptom. It is worth highlighting that participants did *not* require that a measure relate to active symptoms or mimic "real" life to be personally relevant. These findings points to potential differences in participant vs. research conceptions of a "good test," suggesting that a more wholistic evaluation including *context validity* (i.e. the DHT experience from the participant perspective)[24] in conjunction with *content validity* from the research perspective may be needed. A proposed approach to formal assessment of relevance incorporating both of these components is presented in **Figure 3**. Movement towards a more standardized assessment approach that reflects context of use can increase understanding of the relevance of technologies and digital measures across diverse populations, conditions, and monitoring strategies. At the same time, greater transparency in the research process whenever possible (i.e. intentionally educating participants on the purpose and capabilities of new technology) would serve to enhance the relevance of new technologies to people with PD.

Other pragmatic considerations in selection of digital endpoints exist with regards to balancing the psychological impact of testing. Our data present a conundrum: the expressed

desire of people with early PD to proactively monitor for current *and* future symptoms versus the potential for increased anxiety and hypervigilance induced by artificial monitoring. In this study, when participants did well on testing, they felt better about themselves and their disease state. Conversely when they struggled to perform, participants felt more depressed and anxious. Concerns for negative effects of monitoring have been raised in other DHT studies[27] and has been broadly observed with disease screening,[28-31] however it has not been previously reported with remote symptom monitoring in movement disorders. Some participants suggested that retaining "easy" activities helped to offset negative psychological consequences, but it is necessary for scientists to recognize that PD is progressive and decline likely, with potential for increasingly negative psychological impact of testing occurring as disease progresses. Thus, while including "easy" measures may present a short-term solution in early disease, it does not fully address the issue of psychological harm over time. Furthermore, discouragement associated with performing difficult activities (e.g. *Visual-Spatial Memory* task) could impact measurement validity as participants might alter the way that they engage in these tests due to negative impact on self-concept, as was observed in this study. Incorporating measures that allow for a sense of success as well as measures that capture difficulties could improve engagement, perceptions of DHTs, and measurement accuracy by decreasing the individual's use of behaviors that falsely inflate positive outcomes. Alternately, use of unobtrusive passive monitoring approaches may reduce potential for heightened awareness of disease progression and confounding behaviors. Further research is needed in this area to understand the psychological and measurement impacts of active digital monitoring in progressive diseases.

Limitations of this secondary study included use of a predominantly white, college educated and technologically literate study sample. Furthermore, data collection occurred at a single time point in early PD and did not address change over time. Additional research is needed to replicate approaches longitudinally and in more diverse populations. Also, as shown in the results, some people revised initial ratings of relevance over the course of the interview, which was traced back to ambiguity in perceptions of measures. Thus, while strongly correlated, mapping and cognitive interviewing methodologies may offer complimentary perspectives and analysis of discrepancies could support a more nuanced understanding of factors underlying variations in responses, as well as highlighting the urgent need to move beyond rigid binary assessments that are often incompatible with human experience. In total, these findings greatly contribute to understanding the relevance, risks, and benefits of using digital measures to monitor symptoms of early PD from the participant perspective and can inform the use of these tools in clinical trials.

Acknowledgements

The researchers thank the many individuals who contributed to this work. The content is based solely on the perspectives of the authors and do not necessarily represent the official views of the Critical Path Institute, the US FDA or other sponsors. BrainBaseline application screenshots reprinted with permission from Clinical ink.

Author contributions

Authorship contributions were as follows: planning and development of study (all), data collection (JM, JA, PY, MK), data analysis (JA, JM, PY, GS, RS), drafting and revising of manuscript (all).

Disclosures

GTS is an employee of Rush University and has consulting and advisory board membership with honoraria for: Acadia Pharmaceuticals; Adamas Pharmaceuticals, Inc.; Biogen, Inc.; Ceregene, Inc.; CHDI Management, Inc.; the Cleveland Clinic Foundation; Ingenix Pharmaceutical Services (i3 Research); MedGenesis Therapeutix, Inc.; Neurocrine Biosciences, Inc.; Pfizer, Inc.; Tools-4-Patients; Ultragenyx, Inc.; and the Sunshine Care Foundation. He has received grants from and done research for: the National Institutes of Health, the Department of Defense, the Michael J. Fox Foundation for Parkinson's Research, the Dystonia Coalition, CHDI, the Cleveland Clinic Foundation, the International Parkinson and Movement Disorder Society, and CBD Solutions, and has received honoraria from: the International Parkinson and Movement Disorder Society, the American Academy of Neurology, the Michael J. Fox Foundation for Parkinson's Research, the FDA, the National Institutes of Health, and the

Alzheimer's Association. **JC** is Director of Digital Health Strategy at AbbVie and Industry Co-Director of CPP. **TD** Is Executive Medical Director at Biogen. **JH** Is Senior Specialist, Patient Insights at H. Lundbeck A/S, Valby, Denmark. **TS** has served as a consultant for Acadia, Blue Rock Therapeutics, Caraway Therapeutics, Critical Path for Parkinson's Consortium (CPP), Denali, General Electric (GE), Neuroderm, Sanofi, Sinopia, Sunovion, Roche, Takeda, MJFF, Vanqua Bio and Voyager. She served on the ad board for Acadia, Denali, General Electric (GE), Sunovion, Roche. She has served as a member of the scientific advisory board of Caraway Therapeutics, Neuroderm, Sanofi and UCB. She has received research funding from Biogen, Roche, Neuroderm, Sanofi, Sun Pharma, Amneal, Prevail, UCB, NINDS, MJFF, Parkinson's Foundation. **ERD** Has stock ownership in Grand Rounds, an online second opinion service, has received consultancy fees from 23andMe, Abbott, Abbvie, Amwell, Biogen, Clintrex, CuraSen, DeciBio, Denali Therapeutics, GlaxoSmithKline, Grand Rounds, Huntington Study Group, Informa Pharma Consulting, medical-legal services, Mednick Associates, Medopad, Olson Research Group, Origent Data Sciences, Inc., Pear Therapeutics, Prilenia, Roche, Sanofi, Shire, Spark Therapeutics, Sunovion Pharmaceuticals, Voyager Therapeutics, ZS Consulting, honoraria from Alzeimer's Drug Discovery Foundation, American Academy of Neurology, American Neurological Association, California Pacific Medical Center, Excellus BlueCross BlueShield, Food and Drug Administration, MCM Education, The Michael J Fox Foundation, Stanford University, UC Irvine, University of Michigan, and research funding from Abbvie, Acadia Pharmaceuticals, AMC Health, BioSensics, Burroughs Wellcome Fund, Greater Rochester Health Foundation, Huntington Study Group, Michael J. Fox Foundation, National Institutes of Health, Nuredis, Inc., Patient-Centered Outcomes Research Institute, Pfizer, Photopharmics, Roche, Safra

Foundation. **JLA** Has received honoraria from Huntington Study Group, research support from National Institutes of Health, The Michael J Fox Foundation, Biogen, Safra Foundation, Empire Clinical Research Investigator Program, and consultancy fees from VisualDx.

The following authors (JRM, RMS, MLTMM, PY, MC, JEC, SJR, MK, KWB, DS) have no conflict of interest to disclose.

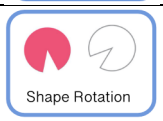
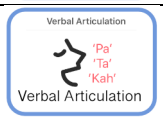
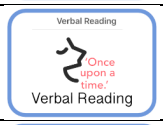
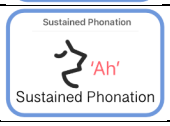
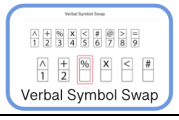
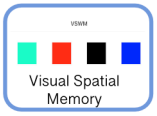
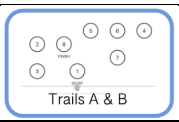
References

1. Smuck M, Odonkor CA, Wilt JK, Schmidt N, Swiernik MA. The emerging clinical role of wearables: factors for successful implementation in healthcare. NPJ Digit Med 2021;4(1):45. DOI: 10.1038/s41746-021-00418-3.
2. Rovini E, Maremmani C, Cavallo F. How Wearable Sensors Can Support Parkinson's Disease Diagnosis and Treatment: A Systematic Review. Front Neurosci 2017;11:555. DOI: 10.3389/fnins.2017.00555.
3. Walton MK, Cappelleri JC, Byrom B, et al. Considerations for development of an evidence dossier to support the use of mobile sensor technology for clinical outcome assessments in clinical trials. Contemp Clin Trials 2020;91:105962. DOI: 10.1016/j.cct.2020.105962.
4. Wilson H, Dashiell-Aje E, Anatchkova M, et al. Beyond study participants: a framework for engaging patients in the selection or development of clinical outcome assessments for evaluating the benefits of treatment in medical product development. Qual Life Res 2018;27(1):5-16. DOI: 10.1007/s11136-017-1577-6.
5. Schultz-Knudsen K, Sabaliauskaite U, Hellsten J, Lassen AB, Morant AV. New Drug and Biologics Approvals in 2019: A Systematic Analysis of Patient Experience Data in FDA Drug Approval Packages and Product Labels. Ther Innov Regul Sci 2021;55(3):503-513. DOI: 10.1007/s43441-020-00244-x.
6. USDHHS. Patient-focused drug development: Methods to identify what is important to patients (Guidance for industry, food and drug administration staff, and other stakeholders). (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-methods-identify-what-important-patients).
7. USDHHS. Patient-focused drug development: Collecting comprehensive and representative input. (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-collecting-comprehensive-and-representative-input).
8. Taylor KI, Staunton H, Lipsmeier F, Nobbs D, Lindemann M. Outcome measures based on digital health technology sensor data: data- and patient-centric approaches. NPJ Digit Med 2020;3:97. DOI: 10.1038/s41746-020-0305-8.
9. Mammen J, Speck R, Stebbins G, et al. Mapping meaningful symptoms and impacts of early Parkinson's disease to digital outcome measures. (Preprint open access) 2022 (https://digitalcommons.uri.edu/nursing_facpubs/343/).
10. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. Value Health 2011;14(8):978-88. DOI: 10.1016/j.jval.2011.06.013.
11. Brain Baseline Ink. 2022.
12. Dorsey E. Wearable Assessments in the Clinic and Home in PD (WATCH-PD). (https://clinicaltrials.gov/ct2/show/NCT03681015).
13. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc 2016;9:211-7. DOI: 10.2147/JMDH.S104807.

14. Mammen J, Norton S, Rhee H, Butz A. New approaches to qualitative interviewing: Development of a card sort technique to understand subjective patterns of symptoms and responses. Int J Nurs Stud 2016;58:90-6. (Research Support, N.I.H., Extramural). DOI: 10.1016/j.ijnurstu.2015.12.011.
15. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Research 2005;15(9):1277-88. DOI: 10.1177/1049732305276687.
16. Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med 2018;18(3):91-93. DOI: 10.1016/j.tjem.2018.08.001.
17. Ayres L, Kavanaugh K, Knafl KA. Within-case and across-case approaches to qualitative data analysis. Qual Health Res 2003;13(6):871-83. (http://www.ncbi.nlm.nih.gov/pubmed/12891720).
18. Saldaña J. The Coding Manual for Qualitative Researchers. 2nd ed. Washington, DC: Sage, 2013.
19. Braun V, Clarke V, Hayfield N, Terry G. Thematic Analysis. In: Liamputtong P, ed. Handbook of Research Methods in Health Social Sciences. Singapore: Springer; 2019.
20. Maxwell JA. Qualitative Research Design: an Interactive Approach. Thousand Oaks, CA: Sage, 2012.
21. Bhidayasiri R, Panyakaew P, Trenkwalder C, et al. Delivering patient-centered care in Parkinson's disease: Challenges and consensus from an international panel. Parkinsonism Relat Disord 2020;72:82-87. DOI: 10.1016/j.parkreldis.2020.02.013.
22. Morel T, Cleanthous S, Andrejack J, et al. Patient Experience in Early-Stage Parkinson's Disease: Using a Mixed Methods Analysis to Identify Which Concepts Are Cardinal for Clinical Trial Outcome Assessment. Neurology and Therapy 2022. DOI: 10.1007/s40120-022-00375-3.
23. Sacks L, Kunkoski E. Digital Health Technology to Measure Drug Efficacy in Clinical Trials for Parkinson's Disease: A Regulatory Perspective. J Parkinsons Dis 2021;11(s1):S111-S115. DOI: 10.3233/JPD-202416.
24. Skinner CH. Contextual validity: Knowing what works is necessary, but not sufficient. The School Psychologist 2013;67(1):14-21.
25. van Halteren AD, Munneke M, Smit E, Thomas S, Bloem BR, Darweesh SKL. Personalized Care Management for Persons with Parkinson's Disease. J Parkinsons Dis 2020;10(s1):S11-S20. DOI: 10.3233/JPD-202126.
26. Bloem BR, Henderson EJ, Dorsey ER, et al. Integrated and patient-centred management of Parkinson's disease: a network model for reshaping chronic neurological care. Lancet Neurol 2020;19(7):623-634. DOI: 10.1016/S1474-4422(20)30064-8.
27. Kanstrup AM, Bertelsen P, Jensen MB. Contradictions in digital health engagement: An activity tracker's ambiguous influence on vulnerable young adults' engagement in own health. Digit Health 2018;4:2055207618775192. DOI: 10.1177/2055207618775192.
28. Oliveri S, Ferrari F, Manfrinati A, Pravettoni G. A Systematic Review of the Psychological Implications of Genetic Testing: A Comparative Analysis Among Cardiovascular, Neurodegenerative and Cancer Diseases. Front Genet 2018;9:624. DOI: 10.3389/fgene.2018.00624.
29. Yang C, Sriranjan V, Abou-Setta AM, Poluha W, Walker JR, Singh H. Anxiety Associated with Colonoscopy and Flexible Sigmoidoscopy: A Systematic Review. Am J Gastroenterol 2018;113(12):1810-1818. DOI: 10.1038/s41395-018-0398-8.

30. Yanes T, Willis AM, Meiser B, Tucker KM, Best M. Psychosocial and behavioral outcomes of genomic testing in cancer: a systematic review. Eur J Hum Genet 2019;27(1):28-35. DOI: 10.1038/s41431-018-0257-5.
31. Gopie JP, Vasen HF, Tibben A. Surveillance for hereditary cancer: does the benefit outweigh the psychological burden?--A systematic review. Crit Rev Oncol Hematol 2012;83(3):329-40. DOI: 10.1016/j.critrevonc.2012.01.004.

Table 1. WATCH-PD digital tasks evaluated for relevance to personally meaningful symptoms from the patient perspective

	TASK	DOMAIN	ASSESSMENT	PICTOGRAPH
Smartwatch	Walking & Balance	Gait/balance	(1) Participant walks straight line for 1 minute. (2) Participant stands with arms at sides for 30 seconds.	
	Tremor Task	Tremor	(1) Participant rests hands in lap for 10 seconds. (2) Participant extends arms out in front for 10 seconds.	
Smartphone Application	Finger Tapping	Fine motor	Participant performs rapid alternating finger movements by tapping two side-by-side targets using index and middle fingers.	
	Shape Rotation	Fine motor	Participant uses 1-2 fingers to move and rotate a pink object into the object outline as quickly as possible.	
	Verbal Articulation	Speech	Participants performs sustained phonation task 15 seconds.	
	Visual Reading	Speech	Participants reads a series of sentences printed on the screen.	
	Sustained Phonation	Speech	Participants repeats the syllables “pa ta ka,” for 15 seconds.	
	Digit Symbol Substitution	Thinking	Participant is presented with a symbol and must speak aloud the corresponding number from a key.	
	Visuo-Spatial Working Memory	Thinking	Participant is briefly shown four different colored boxes followed by a single, colored box and must indicate if the single box was in the previous set of four.	
	Trail Making Task	Thinking	Participant must trace a set of alpha-numeric dots as quickly and accurately as possible using the index finger of the dominant hand.	

Notes. BrainBaseline application screenshots reprinted with permission from Clinical ink.

Table 2. Demographic characteristics of the sample comparative to the parent WATCH-PD study

	Sample n = 40	**Parent study (n = 82)**
Age, years	63.9 (SD 8.8)	63.3 (SD 9.4)
Female, n (%)	19 (47.5%)	36 (43.9%)
Race/ethnicity, n (%)		
White	37 (92.5%)	78 (95.1%)
Asian	3 (7.5%)	3 (3.7%)
Not specified	-	1 (1.2%)
Hispanic or Latino, n (%)	1 (2.5%)	3 (3.7%)
Education > 12 years, n (%)	40 (100.0%)	78 (95.1%)
PD duration, years	2.1 (SD 0.9)	0.8 (SD 0.6)
Taking medications for PD, n (%)	16 (40.0%)	-

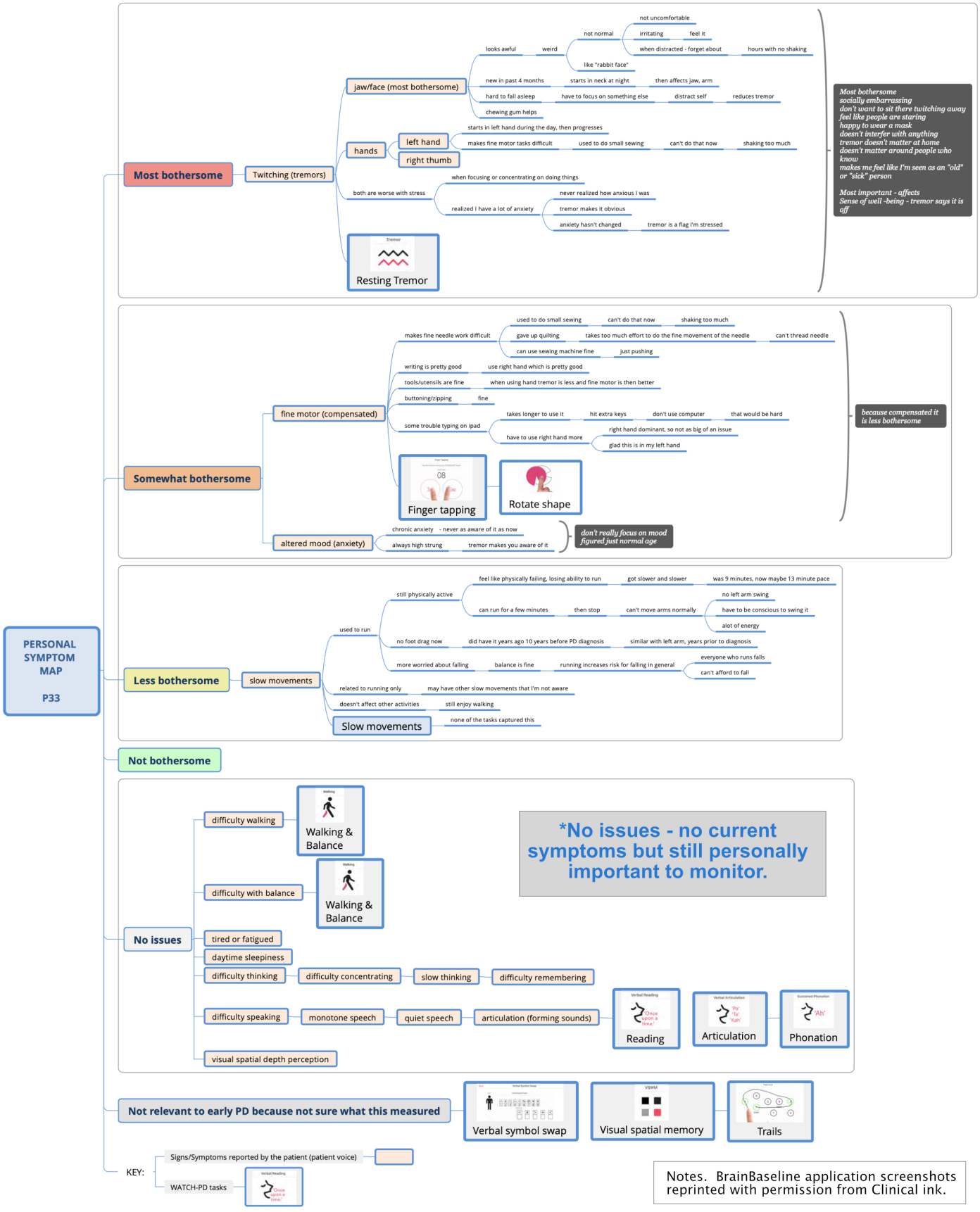


Figure 1. Sample symptom map with pictographs connecting WATCH-PD digital measures to personally important symptoms

Table 3. Frequency and percentage of positive demonstration of understanding and relevance of WATCH-PD tasks during cognitive interviewing vs mapping

WATCH-PD Task	Understood what the WATCH-PD task was asking them to do in order to complete it n/40 (%)= [yes]	WATCH-PD task related to personal PD symptoms n/40 (%)= [yes]	WATCH-PD task related to tasks and activities in daily life n/40 (%)= [yes]	WATCH-PD task was important to measuring PD progression n/40 (%)	WATCH-PD task was relevant to monitoring any personally important PD symptoms	Relevance of WATCH-PD tasks % Agreement A:B	Cramer's V Correlation[c] A:B
	A. COGNITIVE INTERVIEWING				B. MAPPING		
MOVEMENT							
Walking and Balance Task	40 (100%)	26 (65%)	39 (97.5%)	**Yes = 38 (95%)** No = 0 (0%) Qualified[b] = 2 (5%)	**Yes = 37 (92.5%)**	37/40 (92.5%)	0.397**
Resting Tremor Task	40 (100%)	34 (85%)	32 (80%)	**Yes = 39 (97.5%)** No = 1 (2.5%) Qualified = 0	**Yes = 40 (100%)[d]**	39/40 (97.5%)	[d]
FINE COORDINATION							
Finger Tapping Task	39 (97.5%)	34 (85%)	32 (80%)	**Yes = 37 (92.5%)** No = 2 (5%) Qualified = 1 (2.5%)	**Yes = 36 (90%)**	38/40 (95%)	0.604 **
Shape Rotation Task	40 (100%)	35 (87.5%)	34 (85%)	**Yes = 35 (87.5%)** No = 3 (7.5%) Qualified = 2 (5%)	**Yes = 37 (92.5%)**	37/40 (92.5%)	0.589**
SPEECH							
Verbal Reading Task	40 (100%)	15 (37.5%)	36 (90%)	**Yes = 35 (90%)** No = 2 (5%) Qualified = 3 (7.5%)	**Yes = 35 (87.5%)**	34/40 (85%)	.457**
Sustained Phonation Task	40 (100%)	17 (42.5%)	32 (82%)[1]	**Yes = 32 (87.5%)** No = 3 (7.5%) Qualified = 5 (12.5%)	**Yes = 33 (82.5%)**	29/40 (72.5%)	.340*
Verbal Articulation Task	40 (100%)	15 (37.5%)	30 (75%)	**Yes = 27 (67.5%)** No = 5 (12.5%) Qualified = 8 (20%)	**Yes = 34 (85%)**	29/40 (72.5%)	.474**
THINKING							
Trails A&B Task	40 (100%)	23 (57.5%)	25 (62.5%)	**Yes = 34 (85%)** No = 4 (10%) Qualified = 2 (5%)	**Yes = 32 (80%)**	29/40 (72.5%)	0.486**
Verbal Symbol Swap Task	40 (100%)	14 (35%)	33 (82.5%)	**Yes = 33 (82.5%)** No = 1 (2.5%) Qualified = 6 (15%)	**Yes = 33 (82.5%)**	30/40 (75%)	0.301*
Visual/spatial Memory Task	40 (100%)	19 (47.5%)	23 (59%)[a]	**Yes = 21 (52.5%)** No = 12 (30%) Qualified = 7 (17.5%)	**Yes = 21 (52.5%)**	27/40 (67.5%)	.462**

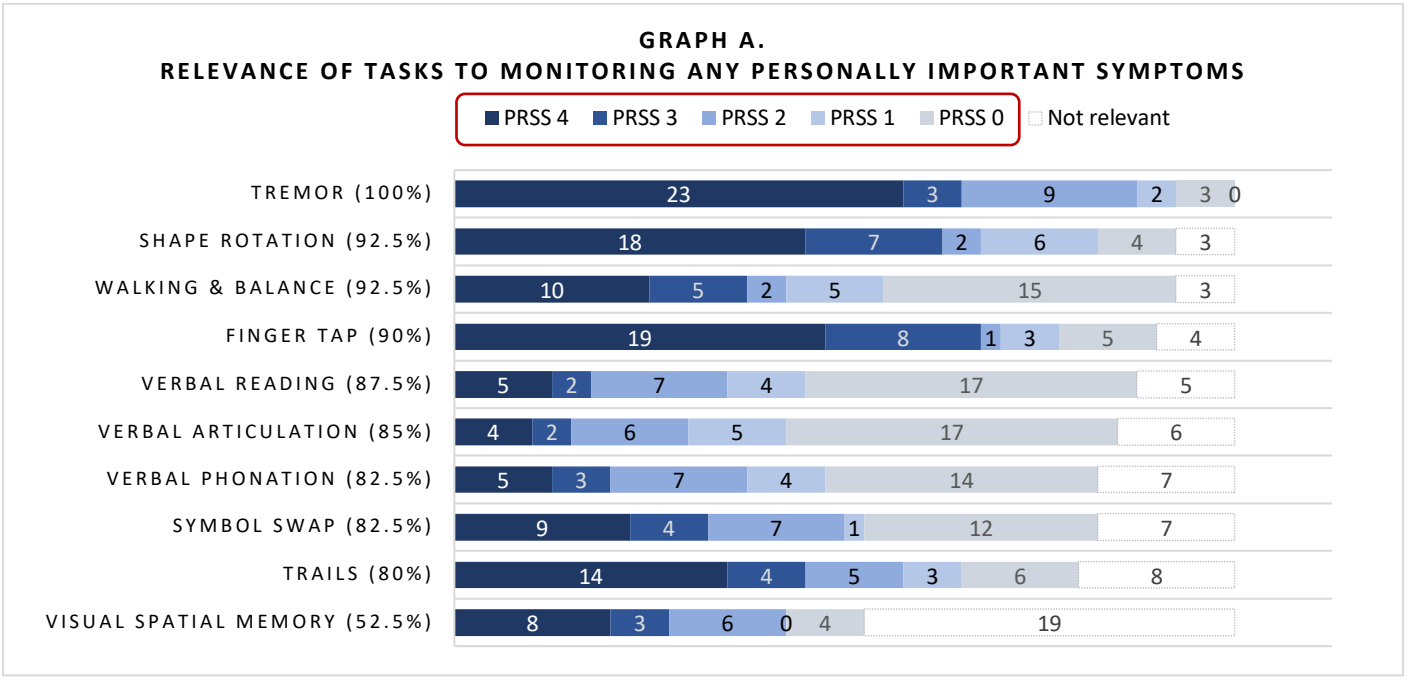
[a] N=39;

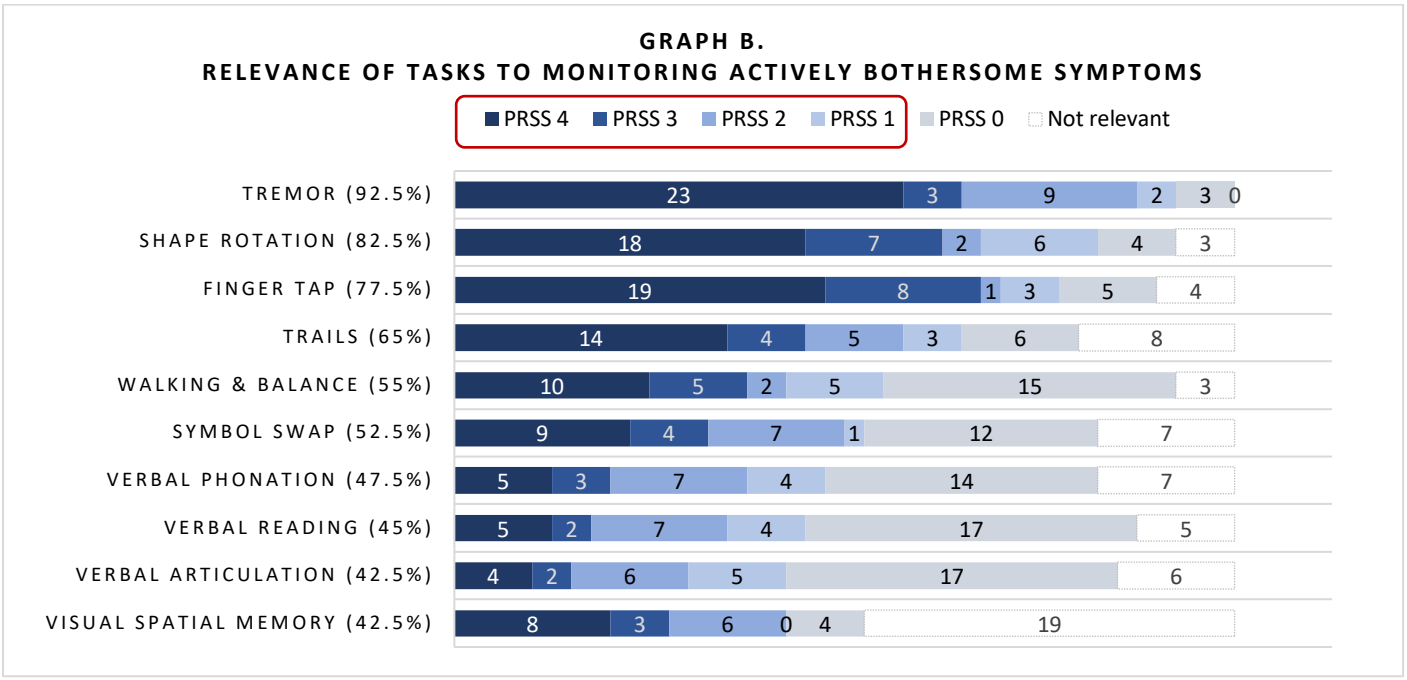
[b] Yes = definitive endorsement of relevance; No = definitive rejection of relevance; Qualified = marginal answer in which participant could not categorically endorse or reject personal relevance and gave a qualified answer (ex. "Yes, *but* only if they fixed the technical issues" or "No, but if I developed symptoms, I would want to monitor for it")

[c] Cramer's V is an effect size of Chi Square test of independence between categorical variables. Polychloric Correlation was not used due to non-normal distribution of data.

[d] Unable to run Cramer V due to lack of homogeneity in response to *Tremor* task on mapping. Cramer's V effect sizes *>0.25 are considered to indicate a strong relationship and **>0.35 a very strong relationship (df=2).

Figure 2. Perceived relevance of WATCH-PD Tasks to monitoring any personally important (experienced or anticipated) versus bothersome symptoms in patients with early PD as elicited by Mapping





Key: Tasks are graded by association with Patient Reported Symptom Scores in symptom maps (PRSS)
Personally important symptoms include PRSS 0 to 4 and may be active or anticipated
PRSS 4=Relating to symptoms that are present and most bothersome;
PRSS 3= Relating to symptoms that are present and somewhat bothersome symptoms;
PRSS 2= Relating to symptoms that are present and less bothersome;
PRSS 1`= Relating to symptoms that are present but not bothersome;
PRSS 0=Relating to symptoms that are not present or bothersome but still perceived as important to monitor over time.
Not relevant = Task perceived as failing to capture any personally important symptoms of PD.
Percentage (%) represents the total percent of participants who viewed the task as relevant to monitoring progression of personally important symptoms of PD.

Table 4. Theme 1 supporting data - Examples of task endorsement based on five criteria used by participants to explain rationales for personal relevance

Understood purpose	Important symptom	Good test of symptom	Related to daily life[a]	Experienced symptom[b]	Endorsed Task[c]	**Interpretation**: Personal relevance of digital tasks to patients for monitoring important aspects of disease is contingent upon: • (C1) approximately understanding what symptom the task is assessing, • (C2) belief that the task is assessing an important symptom of the disease, and • (C3) belief that the task is a good test of that symptom. The task does not have to be related to real life (C4) or an experienced symptom (C5) to be relevant. However, positive endorsement of these criteria can add value to perceived relevance.
C1	C2	C3	C4	C5		**CASES IN WHICH PARTICIPANTS ENDORSED TASK**
✓	✓	✓	✓	✓	✓	**P5**: This was measuring fine motor …. I started to do the rapid tapping, I would lose control of my index finger and the middle finger and they would kind of spazz out. I could tap away a much greater rate with the right hand. [The task relates to my] dexterity, fine motor skills in my right hand. It's very similar to typing. Like I said, this one was the Parkinson's symptom that I'm noticing the most. Yes [it's important to measuring my PD progression]. (*Finger Tapping*)
?	✓	✓	✓	✓	✓	**P25:** I had to walk continuously for a minute and then at standstill for 30 seconds with my hands by my side. I think it was measuring my tremor while I was walking and standing still, since tremor is my only issue. I know it was probably in general looking at gait and balance as well [but I] don't really have those issues and that's why I wasn't focused on them. [The task is important and] similar to just standing still and walking. (*Walking & Balance Task*)
✓	✓	✓	✓	✗	✓	**P24: T**here were four or five sentences that I had to read**.** My understanding is that sometimes people with Parkinson's have difficulty speaking, or their voice will kinda waver out. It'll get quieter. I don't think I have that problem. [but] Your ability to speak and make your needs known is important. [The task is similar to life because] I talk all the time, and if your voice is trialing off you have a problem communicating. [So the task] is very important. I think communication is key. Your ability to verbalize is extremely important. (*Verbal Reading Task*)
✓	✓	✓	✓	✗	✓	**P26:** It was asking me to walk for a full minute back and forth…it was sensoring for tremor and for gait. [I don't have this symptom [but] I do a lot of walking [and] I'm worried. This is one of the worrisome symptoms of PD, for me, not being able to get around, [so the task is important]. I mean I'm watching myself. When I stub my toe, I sit there and I go, was that [PD]? Yeah. I'm already monitoring it. (*Walking & Balance Task*)
✓	✓	✓	✗	✓	✓	**P33:** You sat down, and you put your hands in your lap and then You put your hands out. I thought this was measuring how much you were shaking. {I have] tremor in my hand. I don't sit around with my hands in my lap or my hands out. [But it's important] Because if the watch was picking up the intensity of the tremor, it would show that it was is getting worse. [*Tremor Task*]
✓	✓	✓	✗	✓	✓	**P37:** You keep tapping your index finger and your middle finger over and over, back and forth, both dominant and non-dominant…it's to see how your fine motor skills are working…I didn't realize this until I was doing this, how slow my right hand is compared to my left hand. I might hate this one, but this was one of the most important ones. (*Finger Tapping*)
✓	✓	✓	✗	✓	✓	**P20**: You had to move your fingers to the different numbers quickly…[it was measuring] motor skills, moving from one number to the next and doing it in a quick enough way, and some memory, 'cause you gotta go from one to two to three. [It relates to my difficulty with] motor skills as well as remembering. I think over a period of time, if you struggled with moving from one to two to three and so one, it would be measuring [progression] for sure. (*Trails A&B*)
✓	✓	✓	✗	✗	✓	**P6:** I don't have tremors… I understand why you're testing, and I think it's important to test, but, for my daily life, it wasn't particularly important to me at this point …[it isn't like anything I do in real life, but] I can see where it would be very important. I absolutely believe it's valid, as valid as the walking, yes. [*Tremor Task*]
✓	✓	✓	✗	✗	✓	**P21:** It was asking me to repeat a three-syllable word, pa, ta, ka, pa, ta, ka, pa, ta, ka. [It was measuring] speech and ability to keep syllables separated. I don't think I have any [symptoms] that were related to that task. It's not at all similar [to things I do in daily life]. I have to articulate, but I don't say things over and over and over again. I don't try to squeeze in long sentences in one breath. [But] I think this task would be easy to see when deficits crept in … it would be easy to see a decline with this task. I believe it's important. (*Verbal Articulation Task*)

C1	C2	C3	C4	C5	Endorsed Task	
?	✓	?	✗	✗	✓	**P22:** I was asked to repeat a nonsense word as many times as I could during a period of time. I had to take a deep breath, so that I could say it as many times as possible in that timeframe, so I thought it might have been measuring what my lung capacity was. I don't think I have any symptoms that relate to this task. Generally speaking, I'm not repeating words 15 or 16 times in a row so it's not an example of something that I would do in my everyday life. [But] I guess if you were to look at the data over the course of several months, there might've been trend, so I feel like there might be meaning in the data, so that's what I mean by [I think it is an important task]. (*Verbal Articulation Task*)
C1	**C2**	**C3**	**C4**	**C5**	**CASES IN WHICH PARTICIPANTS DID NOT ENDORSE TASK**	
✓	✓	✗	✓	✓	✗	**P14:** [I had to] see if I can line up the solid circle with the outline circle and how long it would take me… I think[it measured] motor skills and dexterity. The reason I didn't like this one is I didn't think it measured Parkinson's; it just measures general coordination. This would be like trying to screw on a jar; I thought this one was too glitchy with the computer. there were times where my hand was moving, and the thing wasn't moving with me. I didn't think this measured accurately. (*Shape Rotation Task*)
✓	✓	✗	✓	✓	✗	**P26**: It was asking me to take little pie slices and line them up. [It was measuring] my ability to control my hands. [But my] fingernails they get in the way on this exercise. It doesn't let you connect with the pad. To consciously keep that contact, I had to go slower than I normally would. [it's similar to] working on the computer. Yeah, using a mouse, using a track pad. I think that's relevant. [It relates to my trouble with] clicking on a link on your computer, and some of the fine motor skills. [But] I think there's too many things wrong with it for it to be important. Is it the Parkinson's or is it because I'm not positioned properly here? I don't trust this exercise. (*Shape Rotation Task*)
✓	✓	✗	✗	✓	✗	**P20**: [The task measured] cognitive ability to remember colors. [It relates to] my issues remembering and being able to do things quickly. [But] You kinda give up on the task because it's so hard, and sometimes I just guessed… I think it would be challenging for anyone [regardless of PD], and I don't see where you would use it. My fiancée saw it and was like, "I don't think I could do that either." So I would question it's [validity] . (*Visual Spatial Memory Task*)
✓	✓	✗	✗	✗	✗	**P21**: This [task] is where you'd see the four colored boxes and then have to say the one that they flashed up if it was one of four …[It was measuring] Short term memory probably. It was frustrating… I'd get to the end of it and go, "Geez, did I get any of those right?" I don't think [I have] any [symptoms related to it], at least not yet. I don't have to make those kind of decisions that quickly in real life. I'm not sure it was relevant. I felt so crappy about how well I did at it. I don't know if that was related to my Parkinson's or not. I think I would've done crappy at it with Parkinson's or without Parkinson's. It just seems random …I don't know if it showed a progression in my Parkinson's at all. (*Visual Spatial Memory Task*)
✓	✓	✗	✗	✗	✗	**P23: I**t's another fine motor task. You had to put two fingers on the colored circle and twist it and move it until it matched the circle on the inside. I don't have a lot of fine motor issues. I can't think of anything [in life it relates to]. it's not necessarily measuring fine motor skills. It's measuring your ability to learn how to game the system—it's affected by the length of your fingernails, and that doesn't have anything to do with fine motor skills. If it's measuring the length of my fingernails, that's a measurement error. (*Shape Rotation Task*)
?	✗	✗	✗	✗	✗	**P18**: You had to remember these four colors, and then some other color, and match them up. I suppose it's [measuring] your memory. [but] I don't know how it relates to Parkinson's. [I've never had] very strong visualization. I'm not thinking of anything [in real life that it is similar to]. I guess I have no idea [if this is important to monitoring PD], which inclines me to think, no [it's not relevant]. I don't know …I am not expert. (*Visual Spatial Memory Task*)
✗	✗	✗	✗	✗	✗	**P12**: This task, I don't think I'm experiencing any symptoms, and I'm not sure what it was trying to measure. I don't see the connection to [daily life]and I'm not drawing the connection to the Parkinson's piece. (*Verbal Articulation*)
✗	✗	✗	✗	✗	✗	**P24:** You have to look at one, two, three, four [colors], and then the fifth square at the bottom and figure out if that fifth square at the bottom is in any of those four. You know I'm not sure—I'm guessing that it's cognition, [but] … I don't know what this would measure. I don't know [if I have related symptoms] because I don't understand the task. It's not similar to anything in life. I didn't like this. I was unsure of it and I found it extremely frustrating. It's not so much [the difficulty of] the test, it's trying to figure out how does it fit with Parkinson's? How is this going to benefit me, or other people? This was the only test that I recall, that made absolutely no sense whatsoever. I think, ultimately, that's the frustration. (*Visual Spatial Memory Task*)

Notes. Participant criteria for evaluating personal relevance of WATCH-PD tasks to monitoring personally important symptoms of Parkinson's disease.

C1. CRITERIA #1: Believed they knew what task was measuring

C2. CRITERIA #2: Believed the task related to an important symptoms of PD (experienced or not)

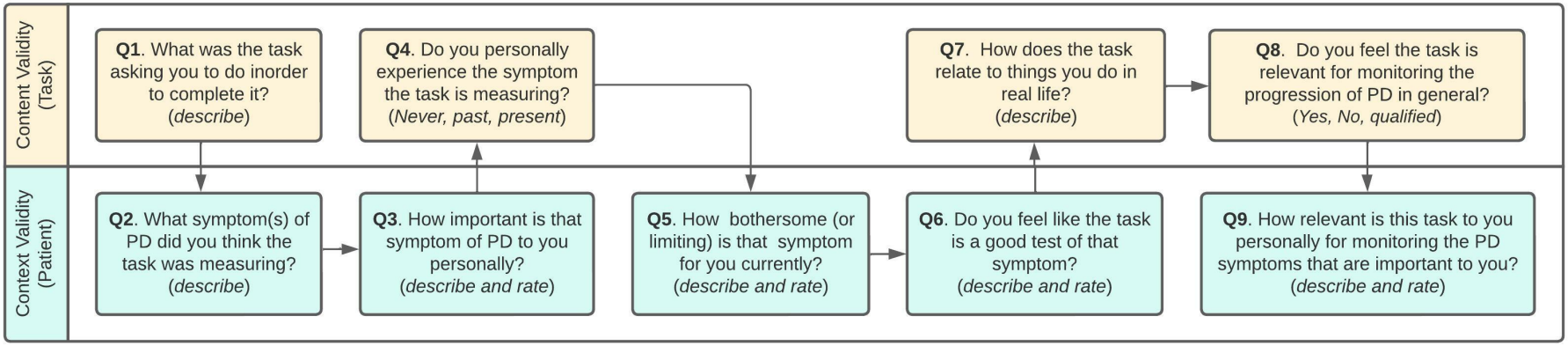
C3. CRITERIA #3: Believed the task was accurate, consistent, and reliable.

C4. CRITERIA #4: Believed task related to activities in performed in real life ([a] Cognitive interview item Q3 – Task relates to activities in real life_

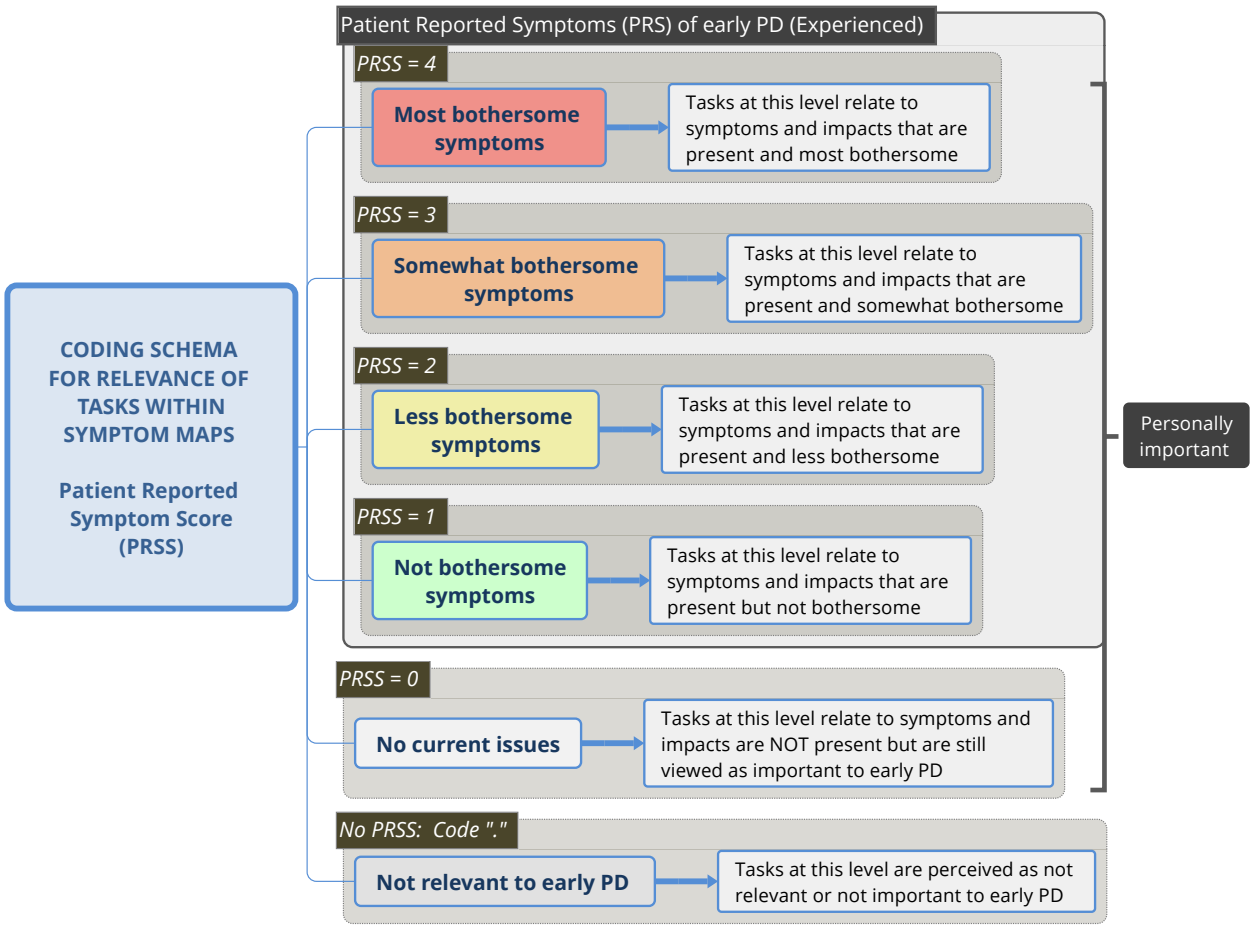
C5. CRITERIA #5: Personally experienced the symptom believed to be measured - past or present ([b] Cognitive interview item Q2 – Task relates to personal symptom of PD)

Endorsed Task = Endorsed task as valuable for monitoring PD symptom progression over time. ([c] Cognitive interview item Q4 – Task is relevant to monitoring personal PD progression)

Figure 3. Recommended approach to assessing the relevance of digital measures for monitoring meaningful symptoms of disease



Notes. Use of a consistent 0-10 rating scale for each rated item (i.e. 0=not important at all; 10=most important, etc.) could improve comparison across technologies and trials.



Supplement A. Coding schema for assessing relevance of digital measures to meaningful symptoms

Supplement B. Supporting data for Theme 2

Theme 2. People with early PD believe it can be important to complete tasks measuring aspects of disease and symptoms that they do not currently experience.

P4: I just know that's something that does impact Parkinson's patients, so I thought that would've been something you could've measured.

P5 All of these things [are important], knowing whether or not there was any noticeable deterioration or whether they stayed the same or improved

P6: I don't recall that I even had any tremors at all. … but I'm curious if it could tell beyond what I can sense. That would be of great interest to me.

P7: Well, it could be important a little longer term…you've got to compare it to something. That's what I'm wondering. It was valuable [to know] if it started and the cognitive issues maybe became more pertinent or speech was slurred or something…. I think with a lot of these tasks that would be something that I would be interested in. Have they changed one way or the other over time or not changed at all over time?

P8: [It's important because] If I had difficulty walking, this [task] would have uncovered that. It would have diagnosed it and sent that information to somebody…or… if I was doing 200 taps, and then the next week only did 100 taps, that would show that something's going wrong, right? In my case, it did not, because not a lot is going on. I think this, again, falls into the general category that if over time, I noticed I was unable to do this, or I was getting worse at it, I would want to know that and use that as part of my discussions with my doctor. I think everything I did fell in that category.

P10: I suspect that this is something that over a longer duration of time would show evidence of deterioration. I think that the articulation issue is an important one and the one I experienced, but not maybe as noticeably at this point in my life. I do think it's important.

P15: I would say yes [it's important], because, over time, it may start to show up. It'll be like, "Okay, six months ago, it didn't tremor…a year later, now it tremors." It doesn't register anything now, but I think over time, if my hands started to tremor, then that would be valuable to know that it's progressing.

P19: I don't have an issue [with speech]. [But] I think the tasks are good to have in the test, because a lot of people with Parkinson's do have these issues. How would I know if I did or didn't have them without doing the task? [and] I'm gonna take back my last response about is this important for measuring Parkinson's, because maybe it is. Just because I didn't have problem with the cognitive doesn't mean that you wouldn't be able to identify issues with other people.

P24: I think it's not so much what you have currently, it's the progression. It's about whether new symptoms develop and if those symptoms become more severe over time. … because you're always thinking at the back of your mind, "How much time do I have? Am I gonna be in a wheelchair? Am I gonna have dementia? Am I gonna be able to take care of myself?"…Symptoms they change, they get worse, or they suddenly show up, and you didn't have it before.

P26: My thinking is that the earlier you catch it, the better. Even if it wasn't a problem [now], I'd want to keep testing it because chances are, it's gonna show up, and we might as well know as early as possible. … just because I don't [have] it now doesn't mean it's not gonna be relevant three months from now. You're going to find me saying yes to almost everything here unless I think it was ill-designed.

Supplement C. Supporting data for Themes 3

Theme 3. Disease monitoring tasks can have negative emotional impact that affects sense of well-being and may impact perception of or engagement with the digital measure.

P1: Whenever I take a survey—they go through pages and pages of, "Have you choked yet? Have you fallen yet?" all these horrible things, "Are you constipated?" it's like, I don't have any of those things, but should I, will I? Will I have all of them? That's depressing

P5: I like to feel like I'm doing well, when I'm testing on these things in this way, I'm pretty clearly, I felt like it was not.

P6: I think in general, I'm mostly on top of things, but this I wasn't on top of. I just couldn't do it that fast…I wasn't going to let it stress me out. I just go, "Oh, I hate this. I know I'm not gonna do well." As hard as I tried to do well, I just never felt like I was on top of it.

P10: it made me aware of issues I didn't really know I had. I was blaming the game, not my cognitive difficulties, but it's a stupid game. It illustrated my problems with remembering. It's just right now that I'm remembering how stressful it was.

P11: I think I notice when I do this test that I have variations in tone and amplitude that I'm not in control of. I don't know if that's just everybody or if it's PD related, or if it's just me. It makes me conscious that I have less control in an area that I'm not really aware of.

P14: I hated doing. I guess it pointed out to me how slow my hand is getting. I was like, "Why can't I do this?" …. This one always made me feel really stupid. I never felt like I was doing it good. I think it was a tricky test, but I think it would measure whether your cognitive thinking, your ability to retain information and spit it back out is getting slower over time. I just didn't like it 'cause it made me feel dumb.

P15: That belongs under measure of insanity. I know I did very poorly on all the way through from the beginning.

P16: This one caused me the most anxiety. I bet that if you looked at the accuracy at the beginning of the test, and compare it with the accuracy at the end of the test, that the percentage correct answers would be far better in the beginning of the test than at the end. By the end, people are just so frustrated that they'll tap any finger. They won't care anymore.

P17: I felt like it was beating me. It's like playing poker with the house, the house wins every time. I felt like I did not have the cognitive ability to beat the house…. I felt like I was really bad at it.

P19: I didn't like being asked those questions every day … Instead of just going about my day [happily] It made me feel moody. …. I was trying to do my best, but I was like, my best ain't good enough for this one.

P21: It was frustrating in the sense that I didn't feel like I was very good at it. I don't want to see the results I would get in that.

P23: A big issue with this, is you bring in previous frustration and it affects you even more. In other words, as frustration increases, the failure rate on the test is going to increase. Over time, you may be measuring more the frustration level than the Parkinson's level.

P26: when the doctor's testing you, you're really there to perform and so you get anxious. I felt like it was rigged against me. I felt like oh no, I can't handle this. This is gonna make me feel so inadequate and I'm gonna get depressed. ..After I was all done, I would go be depressed for an hour because it pointed out exactly what was going on—it pointed it out to me that I was having trouble with some of these things.

P27: I completed the test, but I dreaded it, and I'm like, "Can't I just skip this?" and I'm thinking, "No, you've gotta finish it all." But it made me anxious…. "Oh, my gosh. This is gonna be awful, and they're gonna think I've progressed." I just wanted to say, "Oh, the heck with this," {Someone] mentioned that this was hated by all of the participants, I felt better. I thought, "Okay, it's just not me. Everybody's screwin' up the test."

P33: Things that I could do when I started it, I couldn't do with my left hand… it was a reminder I was getting worse. When we did the study, they would ask you all these questions about symptoms, it'd be be depressing where you could be going with this. It would take me a day or two to forget about it.

P36: [it made me] little bit anxious approaching the test. Wondering what my numbers are gonna be, if it's getting worse.

P37: It was absolutely horrible. I didn't realize until I was doing this how bad my right hand is compared to my left hand….Is it getting worse? What do we need? Do we need to go on medication?" It's 100 things running through my brain.