OPEN

Alternatives for Measuring Sitting Accumulation in Workplace Surveys

Bronwyn K. Clark, PhD, Samantha K. Stephens, PhD, Ana D. Goode, PhD, Genevieve N. Healy, PhD, and Elisabeth A.H. Winkler, PhD

Objectives: To develop and assess the measurement properties of self-report measures of accumulation of sitting time. **Methods:** Seven candidate measures were collected in 51 workers from three office environments (79% women) via online questionnaire administered immediately before and after 7-day monitoring periods (activPAL3 24-hour protocol with diary recorded work hours). **Results:** Three measures had some validity ($P < 0.05$ vs activPAL): % of sitting in long bouts more than or equal to 30 minutes, sitting strategy frequency (0 to 100), and interruption rate (n/h sitting). Agreement was limited. Some reliability (intraclass correlation or kappa $P < 0.05$) was seen for these measures, strategy variety (0 to 100), typical day (five categories), and making a conscious effort to sit less (yes/no). **Conclusions:** Two brief and one longer option may suit workplace studies requiring self-report measures of sitting accumulation. Validity was weaker for sitting accumulation than sitting time.

Keywords: accumulation, occupations, office, sitting position, surveys and questionnaires, work

Public health research has identified excessive sedentary behavior as a risk factor for chronic disease and premature mortality.[1] Sedentary behavior is defined as any waking behavior characterized by energy expenditure of less than or equal to 1.5 metabolic equivalents, while in a sitting, reclining, or lying posture.[2] Engaging in large quantities of sedentary behavior has been associated with increased risk of noncommunicable diseases (type two diabetes, cardiovascular disease, and some cancers), musculoskeletal disorders, and mental health disorders.[3–5] Furthermore, accumulation of uninterrupted sedentary behavior for prolonged periods at a time has been shown to confer particular cardiometabolic risk.[6] Thus both amount of sitting time and how it is accumulated are important. For this reason, physical activity guidelines recommend reducing overall sitting time and breaking up periods of prolonged sitting.[7,8]

Desk based workers may be particularly exposed to long periods of sitting as it is estimated that an average of 75% of time spent in an office work environment and 90% of call center work time is spent sitting, much of which is accumulated in prolonged unbroken bouts.[9–11] Accordingly, such workplaces have become prime targets for interventions to reduce and break up sitting time.[12] Appropriate measurement methods are required to monitor sedentary behavior and evaluate its change over time and in interventions: both the amount of sitting time and how it is accumulated. Some activity monitors, such as the commonly used activPAL device,[13] have been shown to accurately measure both amount and accumulation of sedentary time.[14] However, their expense and logistical requirements, including technical expertise, are sometimes prohibitive for such research grade devices. Questionnaire measures are comparably inexpensive and accessible, reaching a wider population more affordably,[15] but they have the potential for error and recall bias, and thus require testing of their measurement qualities (eg, reliability, validity) to provide evidence of how fit for purpose they may be.

By contrast with the numerous questionnaire measures targeting the amount of sitting time[16] there are very few questionnaire measures of sitting time accumulation in the literature.[16] Notably, none are published as having high validity.[16] To measure sitting time accumulation, questionnaires predominantly ask participants to recall how many breaks they have taken per hour of sitting, resulting in measures with low correlations against objective criteria (Spearman correlations between 0.02 and 0.39).[16–19] Questionnaires seldom ask about other aspects of sitting time accumulation, such as how long at a time participants sit or the context in which sitting is occurring and then interrupted.

In view of the limited existing options for workplace studies, and their poor validity, we developed and tested a range of self-report measures of sitting time accumulation. The range was deliberately diverse, including shorter and more detailed alternatives, as well as questions that enquire directly about accumulation, and about relevant behaviors, behavioral intentions and general impressions that may indirectly capture sitting accumulation patterns. We specifically reported on reliability, minimal detectable change (MDC), and validity relative to accurate measures from the activPAL, and accordingly provided guidance on the suitability of the self-report measures for different types of studies. To provide further context, the measurement properties tested for sitting accumulation were also reported for workplace sitting time, standing time, and moving time.

From the School of Public Health, The University of Queensland, Brisbane, Australia.

Sources of Funding: The BeUpstanding program was supported by funding from the Queensland government ''Healthier. Happier. Workplaces'' program, Safe Work Australia, Comcare and the National Health and Medical Research Council of Australia (NHMRC) through a Partnership Project Grant (#1149936) conducted in partnership with Comcare, Safe Work Australia, the Queensland Office of Industrial Relations, VicHealth, and Healthier Workplace WA. BKC is supported for salary by an NHMRC of Australia Early Career Fellowship (#1107168). SKS was supported by an NHMRC Partnership Project Grant (#1149936). GNH was supported by a Medical Research Future Fund MRFF-NHMRC Investigator Grant. NHMRC had no role in the study in terms of the design, data collection, management, analysis and interpretation.

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Approval of Studies/Informed Consent: Ethics approval was obtained from the University of Queensland Human Research Ethics Approval Committee (approval number 2016001743). Participants provided informed online consent.

Statement of significance to clinical practice: This paper reports the measurement qualities of seven candidate self-report measures of the accumulation of sitting time, with corresponding recommendations and contraindications for their use in different study designs. Two measures (one brief, one longer) were both valid and reliable. A further measure was valid but had limited reliability.

Supplemental digital contents are available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.joem.org).

Address correspondence to: Bronwyn K. Clark, PhD, School of Public Health, The University of Queensland, Herston Rd, Herston 4006, Australia (b.clark3@uq.edu.au).

DOI: 10.1097/JOM.0000000000002387

METHODS

Recruitment for this study was conducted as part of BeUpstanding—an online workplace health program aimed at reducing and breaking up sitting time that is being evaluated in the context of a national implementation trial in Australia.[20] Brisbane based workplaces with teams participating in the program were considered eligible for this sub study. Three teams expressed interest and agreed for their staff to participate. The first workplace consisted of mostly call center workers who were part of a large health advice organization. The second and third workplaces involved general office based work and were project management (small business) and government (medium business) organizations. Staff in the team were informed about the study and invited to participate via an email sent by the workplace contact for the BeUpstanding program. The email included an online link to the consent page. Once participants consented, and confirmed their eligibility (desk based workers, ambulatory) they were directed to the first survey.

This survey (Time 1) contained demographic questions (sex, age, education) and the sitting time and accumulation questionnaire. Participants then wore an activity monitor (activPAL3 micro) and recorded wake/sleep times and work start and finish times in an excel based diary for 7 days. At the end of monitor wear, they completed a second administration of the sitting time and accumulation questionnaire (Time 2). Ethics approval was obtained from the University of Queensland Human Research Ethics Approval Committee (approval number 2016001743).

The sitting time and accumulation questionnaire covers a recall period of the last 7 days, and mostly enquired about a typical or average workday over that timeframe. It included the validated Occupational Sitting and Physical Activity Questionnaire (OSPAQ) which asks participants what percentage of work hours are spent sitting, standing, walking, and in heavy labor.[21] Given the paucity of heavy labor in desk based workers in office and call center settings, walking, and heavy labor were considered combined as ''moving.'' The accumulation questionnaire included one previously tested measure regarding breaks,[19] and a range of new measures that were developed by the research team based on existing methods of measurement[21,22] and evidence from interventions regarding the behaviors people perform to break up sitting time.[23] Feedback on the draft questionnaire was sought from three expert peers with modifications made based on their responses. Firstly, a question relating to people being aware of the number of breaks they took during the workday was suggested and a question about conscious effort to break up sitting time was added in response. Secondly, problems with completion and scoring of the open text response format to the original behavioral strategies question were raised, so the revised version uses a predetermined list and asks about whether and how often per day they are used. Finally, the word ''breaks'' had the connotation of resting or not working (eg, lunch break) and was replaced with the term ''interruption'' making specific reference to alternating postures.

The final questionnaire can be found in Supplemental Digital Content (Text, Supplemental Digital Content 1, http://links.lww.com/JOM/A990, copy of questionnaire). Briefly it contained seven candidate measures of sitting accumulation, some that enquire directly regarding accumulation (1 to 3) and some (4 to 7) that may capture accumulation indirectly while enquiring about other facets of sedentary behavior. These were:

1. % sitting in long bouts—This question asked about the percentage of sitting time at work that occurs in bouts of 30 minutes or longer (numeric response). It was designed to immediately follow the OSPAQ.
2. Longest sitting bout–What is the longest period of time you spent sitting down without leaving your seat during the average workday in the last 7 days? (response options: less than 30 min/30 to less than 60 min/60 to less than 90 min/90 to less than 150 min/more than or equal to 150 min).
3. Interruption rate—How many times did you interrupt (by standing or moving) your sitting each hour during an average working day over the last 7 days? (0/1/2/3/4/5 or more). The wording used was aimed to reduce confusion from the term ''break'' in the previously published version.[19]
4. Conscious effort—Did you make conscious efforts to reduce the time you spend sitting behind your desk during the day? (Yes/No).
5. Typical day—How would you describe your typical day at work over the last 7 days? (Mostly sitting with very little standing and moving/mostly standing/changing from sitting to standing throughout the day/constantly moving). Categorical descriptions of usual day occupational and domestic activity have been used previously in the National Health and Nutrition Examination Survey, although with different wording and focus.[24]
6. Sitting strategy variety, and 7. Sitting strategy frequency. These two measures were constructed from a 10-item array in which participants were asked how many times (interval choice of 0 to 10 or more times per day) they used each of a list of 10 strategies to reduce and break up sitting time. The strategies selected for the list were those that were commonly used by workers in previous interventions.[23] Sitting strategy variety was calculated as the number of strategies participants indicated they undertook at least once per day (0 to 10 strategies). Sitting strategy frequency was the sum total of the number of times per day each of the 10 strategies was used (0 to 100 times per day).

Monitor-derived criteria for the questionnaire measures were obtained from the activPAL—which records acceleration in three axes and outputs each instance the wearer spent sitting, standing, or stepping[25]—and a concurrent Excel based diary capturing work start and finish times, wake and sleep times as well as any times the monitor was removed. The activPAL3 micro monitors were initialized with PAL software (version 7.2.32) then waterproofed using nitrile finger cots and a layer of Opsite. Packs containing the waterproofed activPAL monitors, dressings (Hypafix), and instructions were distributed through the workplaces as per recommended practice in field based monitoring using activPAL devices.[26] Based on the written and video instructions, which have been shown to achieve appropriate monitor wear placement,[27] participants self-attached the devices on the midline of the right thigh, one-third of the way down from the thigh crease to the knee.[27] Participants wore the monitors 24 hours per day for 7 days (ie, over the recall period for the second administration of the questionnaire) then returned the monitors to the workplace whereupon they were collected by the research team.

A bespoke SAS (SAS Analytics and Software Solutions, Cary, NC) program was used to combine data from the ''events'' files and the diary and extract a range of summary variables limited to the times the participant was awake, wearing the monitor and working. The amount of work time (min/day, %) spent in each activity (sitting, standing, and stepping) was calculated along with a range of sitting accumulation indicators: percentage of workplace sitting accumulated in bouts of more than or equal to 30 minutes (% sitting in long bouts); longest sitting bout during work (minutes); sit-upright transitions per hour of sitting during work (n/h workplace sitting time); and usual sitting bout duration (minutes). Most variables were calculated as totals per workday then averaged across valid workdays (ie, worn more than or equal to 80% of work hours). Longest sitting bout duration was the observed maximum overall. Usual bout duration—half of all sitting time is accumulated in bouts of this duration or longer—was calculated using non-linear regression.[28] Variables from the activPAL that served as validity criteria were % sitting in long bouts and longest

bout duration (for their self-report counterparts), sit-upright transitions per hour of sitting (for interruption rate, sitting strategy variety, and sitting strategy frequency); and all of the accumulation measures (categorical self-report measures). The OSPAQ % of work hours sitting, standing, and moving were compared against corresponding % of work hours from the activPAL.

Statistical Analyses

Test-retest reliability was assessed using Intraclass Correlation Coefficients (ICC) for agreement (single measures two-way mixed effects models) for continuous variables and using Kappa for categorical variables. ICC values were interpreted as: less than 0.50 (poor reliability); 0.50 to less than 0.75 (moderate reliability); 0.75 to 0.90 (good reliability); and more than 0.90 (excellent reliability).[29] Kappa was interpreted using Landis and Koch description: less than or equal to 0 (no agreement); 0.01 to 0.20 (none to slight); 0.21 to 0.40 (fair); 0.41 to 0.60 (moderate); 0.61 to 0.80; (substantial); and, 0.81 to 1.00 (almost perfect agreement).[30] Related to reliability, for continuous and interval measures, minimal detectable change (MDC) and minimal relative detectable change (MDC%) were calculated, to respectively indicate the absolute and relative magnitude of change that would be greater than the measurement error of the measures.[31] The MDC was calculated at a 90% confidence interval using the equation $MDC = SEM \times 1.65 \times \sqrt{2}$, where SEM is the standard error of measurement, 1.65 is the z-score at the 90% confidence level and the square root of two accounts for errors associated with the repeated measurement. Standard deviation of the measure (difference in the Time 1 and Time 2 scores; sd) and the reliability coefficient (ICC for the group; r) were used to calculate SEM[32,33] using the equation $SEM = sd \times \sqrt{(1-r)}$. MDC% was calculated as (MDC/mean) × 100 where ''mean'' is the mean for all observations in the measure from both test and retest.

Agreement with the criterion was reported for self-report measures that had the same response scale as their criteria as mean differences with 95% limits of agreement. Correlation with the criterion (Spearman correlations, ρ) were reported for all continuous and interval responses. Concurrent validity for the binary measure (conscious effort) and the ordinal measures (longest bout duration and typical day) were respectively tested for their capacity to detect differences between ''yes'' ''no'' and categories via t tests and ANOVA, and trends across ordinal responses in the criterion accumulation measures via the Jonckheere-Terpstra test for trend.[34] Means were also reported to display the magnitude of differences and trends.

Statistical analyses were conducted with SPSS software version 27. Analyses were not stratified on workplace due to the small sample size. Significance was set at $P < 0.05$/95% confidence and at 90% confidence for MDC. All evaluable data were used. The minimum requirements for reliability and MDC were data from both questionnaire administrations, which were approximately 1 to 2 weeks apart, with a median (25th, 75th percentile) of 9.5 (8, 14 days) days between administrations. The minimum requirement for validity was data from the second questionnaire administration plus at least one valid day of activPAL work hours data, with a median (25th, 75th percentile) of 4 (3, 5 days) valid days of work data obtained.

RESULTS

A total of 56 participants consented to take part in the study (workplace 1: 34, workplace 2: 8, workplace 3: 14). Of these, 49 completed surveys at both time 1 and 2 (reliability sample) and 42 completed the survey at time 2 and had valid activPAL data (validity sample). The majority were women ($n = 44$, 79%) and had completed university level education ($n = 54$, 80%). Work hours reported in the diary showed that 19 participants work at least some shift-work hours (ie, outside of 7 am to 6 pm). According to the activPAL, participants sat on average nearly 60% of their workday, with nearly half of this accrued in bouts of 30 minutes or longer. The activPAL revealed that the average sit-stand transitions per hour of sitting was six, the longest bout of sitting was 90 minutes on average, and usual bout duration averaged 30 minutes (See Table, Supplemental Digital Content 2, http://links.lww.com/JOM/A991, sample characteristics). The overall sample characteristics were not necessarily reflective of each workplace (see Supplemental Digital Content 2, http://links.lww.com/JOM/A991), where some variation was seen in: age (workplace 1 oldest); sex (workplace 2 lowest % female); work hours (workplace 1 lowest % working full-time); and activPAL measured workplace sitting time and sitting accumulation (workplace 3 lowest sitting % and % sitting in bouts more than or equal to 30 minutes).

Test-retest reliability according to ICCs was good for sitting %, standing %, and sitting strategy frequency; moderate for moving %, % sitting in long bouts and strategy variety; and poor for interruption rate (Table 1). Mean differences between test and retest were small (≈2% of the mean or less, Table 1) but with wide limits of agreement. There was no agreement between test and retest values of longest bout (Kappa: 0.02, $P = 0.791$); fair agreement for conscious effort (Kappa: 0.39, $P = 0.007$); and moderate agreement for typical day (Kappa: 0.41, $P < 0.001$) (Table 2).

TABLE 1. Test-retest Reliability Over Approximately 1 to 2 Weeks and Minimal Detectable Change ($n = 49$)

	Mean(SD)		Time 2 Versus Time 1		Minimal Detectable Change	
Measure	**Time 1**	**Time 2**	**MD(95% LoA)**	**ICC(95%CI)**	**MDC**	**MDC%**
OSPAQ						
Sitting %	60.3 (22.1)	58.3 (19.4)	2.0 (−28.7, 24.7)	0.79 (0.65, 0.87)	14.7	24.8
Standing %	25.1 (18.3)	26.6 (17.2)	−1.5 (−20.2, 23.3)	0.81 (0.68, 0.89)	11.5	44.4
Moving %	14.6 (9.5)	15.1 (8.9)	0.5 (−16.7, 17.8)	0.54 (0.30, 0.71)	14.0	94.4
Sitting accumulation						
% of sitting in periods ≥30 min, %	55.3 (24.1)	55.5 (24.2)	−0.2 (−39.1, 39.5)	0.66 (0.46, 0.79)	27.5	49.6
Interruption rate, n/h of sitting	2.9 (1.8)	2.7 (1.6)	0.2 (−4.7, 4.4)	0.05 (−0.24, 0.32)	5.3	186.3
Sitting strategy frequency (0–100)*	27.2 (17.4)	26.6 (17.2)	0.6 (−26.4, 25.2)	0.71 (0.54, 0.83)	16.5	61.4
Sitting strategy variety (0–10)	6.9 (2.2)	6.4 (2.2)	0.5 (−3.6, 2.7)	0.73 (0.57, 0.84)	2.3	34.5

ICC, intraclass correlation coefficient; MD (95% LoA): mean difference (95% limit of agreement); MDC, minimal detectable change; MDC%: minimal relative detectable change (percent of the overall mean across Time 1 and 2); OSPAQ, Occupational Sitting and Physical Activity Questionnaire.

*$n = 48$ due to item missing data.

TABLE 2. Test-retest Reliability of Sitting Accumulation Measures With Categorical Responses ($n = 49$)

Measure	*n* (%)		Kappa. *P* Value
	Time 1	Time 2	
Longest sitting bout			0.02, $P = 0.791$
<30 min	1 (2.0%)	1 (2.0%)	
30 to <60 min	11 (22.4%)	10 (20.4%)	
60 to <90 min	15 (30.6%)	18 (36.7%)	
90 to <120 min	14 (28.6%)	11 (22.4%)	
120 to <150 min	8 (16.3%)	8 (16.3%)	
≥150 min	0 (0.0%)	1 (2.0%)	
Conscious effort to reduce sitting*			0.39, $P = 0.007$
Yes	28 (58.3%)	30 (61.2%)	
No	20 (41.7%)	19 (38.8%)	
Typical day			0.41, $P = <0.001$
Mostly sitting	27 (55.1%)	19 (38.8%)	
Mostly standing	3 (6.1%)	5 (10.2%)	
Mix of sitting and standing	19 (38.8%)	24 (49.0%)	
Mostly moving	0 (0.0%)	1 (2.0%)	

*Missing $n = 1$ Time 1.

The MDC for the continuous and interval measures were large and highly varied (Table 1), and larger for sitting accumulation measures than for sitting time (MDC% = 25% for sitting %). The accumulation measure with the best potential to detect change was strategy variety (MDC% = 34%). The poorest potential to detect change was interruption rate (MDC% = 186%), while changes of approximately 50% were detectable for % sitting in long bouts and 60% for sitting strategy frequency.

Criterion validity results are shown in Table 3 for continuous measures and Table 4 for categorical measures. Correlations (ρ) with the activPAL were moderate for % sitting in long bouts, interruption rate, and strategy frequency but weak and non-significant for strategy variety. The strongest correlation, for interruption rate, was similar to those seen for moving % and slightly lower than for those seen for sitting % and standing % (Table 3). The % sitting in long bouts measure showed a modest mean difference against the criterion (–6%) with wide 95% agreement limits (–59%, 47%) (Table 3; Fig. 1) as did the OSPAQ measures (Table 3). The conscious effort question was not significantly related to activPAL sitting accumulation measures (Table 4), but there was a slight tendency for those who made no conscious effort to reduce sitting to have longer sitting accumulation on all the indicators. No significant trends were detected (Jonckheere–Terpstra trend test Table 4) in activPAL assessed accumulation across ordinal categories of the longest sitting bout ($P = 0.436$ to 0.535) and typical day ($P = 0.051$ to 0.184) measures. For typical day, this may have been more about a lack of ordinality across the categories, rather than a lack of difference, with the ''mostly sitting'' group consistently tending to show more prolonged accumulation, longer usual bout duration and maximum sitting bout than mostly standing ($P = 0.023$, 0.019, and 0.016 respectively for associations overall). Changing the order of the categories in this measure to more closely align with the expected amount of sitting time (mostly sitting with very little standing and moving/changing from sitting to standing throughout the day/mostly standing/constantly moving) resulted in significant trends for all accumulation outcomes: %Sitting in periods more than or equal to 30 minutes $T_{JT} = 145$, $z = -2.84$, $P = 0.005$; sit-stand transition rate $T_{JT} = 380$, $z = 2.79$, $P = 0.005$; usual bout duration $T_{JT} = 144$, $z = -2.86$, $P = 0.004$; maximum sitting bout duration $T_{JT} = 147$, $z = -2.79$, $P = 0.005$. For the longest sitting bout question, usual bout duration showed a significant association overall ($P = 0.004$), with longer usual bout for those reporting the highest category of longest bout (more than or equal to 150 minutes); however, only two participants selected this

TABLE 3. Comparison of Self-Report Measures of Sitting Accumulation and Sitting Time With activPAL ($n = 42$)

Measure	Mean (SD)		MD (95% LoA)	Spearman Correlation ρ (95% CI)
	Self-report	activPAL	activPAL - Self-report	
OSPAQ				
Sitting %	58.1 (20.8)	59.1 (22.8)	1.0 (–32.5, 34.6)	0.70 (0.46, 0.88)
Standing %	27.1 (18.8)	33.8 (22.0)	6.8 (–25.5, 39.0)	0.73 (0.52, 0.86)
Moving %*	14.8 (8.7)	7.0 (3.3)	–7.8 (–22.9, 7.3)	0.53 (0.33, 0.69)
Sitting accumulation				
% of sitting in periods ≥30 min, %	55.1 (26.2)	49.6 (21.0)	–5.8 (–57.7, 46.0)	0.42 (0.09, 0.63)
Interruption rate, n/h sitting†	2.7 (1.6)	5.6 (3.5)	–	0.48 (0.23, 0.66)
Sitting strategy frequency (0–100)†	26.5 (17.9)	5.6 (3.5)	–	0.34 (0.04, 0.59)
Sitting strategy variety (0–10)†	6.4 (2.3)	5.6 (3.5)	–	0.21 (–0.10, 0.49)

MD (95% LoA), mean difference (95% limit of agreement); OSPAQ, Occupational Sitting and Physical Activity Questionnaire.
*Walking plus heavy labor (self-report) or stepping (activPAL).
†activPAL criterion measure is sit-stand transition rate (n/h sitting).

TABLE 4. Mean Criterion Sitting Accumulation (activPAL) by Self-reported Patterns of Sitting Time Accumulation

		activPAL Measure Mean (SD)			
Self-report Measure	***n***	**% Sitting in Periods ≥30 min**	**Sit-stand Transition rate n/h Sitting**	**Usual Bout Duration, min**	**Maximum Sitting Bout Duration, min**
Conscious effort to reduce sitting					
Yes	25	47.2 (19.2)	6.0 (3.9)	27.5 (11.5)	83.3 (35.6)
No	17	53.2 (23.5)	5.0 (2.9)	35.2 (21.9)	101.9 (45.2)
t test		t (40) = 0.90, P = 0.372	t (40) = −0.87, P = 0.388	t (40) = 1.50, P = 0.143	t (40) = 1.49, P = 0.144
Typical day					
Mostly sitting	18	58.8 (21.1)	4.8 (4.2)	38.4 (21.0)	107.2 (44.8)
Mostly standing	5	33.6 (12.8)	8.7 (3.7)	19.0 (5.4)	52.3 (17.0)
Mix of sitting and standing	19	45.1 (19.3)	5.5 (2.2)	26.4 (9.7)	85.5 (32.0)
Mostly moving	0	–	–	–	–
Trend		T_{JT} = 182, z = −1.95, P = 0.051	T_{JT} = 331, z = −1.62, P = 0.106	T_{JT} = 189, z = −1.78, P = 0.075	T_{JT} = 208, z = −1.33, P = 0.184
$F_{(df)}$, P		$F_{(2, 39)}$ = 4.19, P = 0.023	$F_{(2, 39)}$ = 2.71, P = 0.079	$F_{(2, 39)}$ = 4.39, P = 0.019	$F_{(2, 39)}$ = 4.63, P = 0.016
Longest sitting bout					
<30 min	1	71.5 (–)	2.7 (–)	45.05 (–)	128.4 (–)
30 to <60 min	11	45.0 (21.8)	7.1 (5.2)	27.1 (17.2)	78.4 (42.8)
60 to <90 min	14	51.3 (17.6)	5.3 (2.9)	30.3 (11.0)	91.1 (35.2)
90 to <120 min	9	44.4 (17.5)	5.2 (1.8)	25.5 (7.7)	91.3 (20.5)
120 to <150 min	5	45.4 (27.9)	5.8 (2.9)	28.9 (19.8)	79.8 (39.0)
≥150 min	2	86.2 (2.0)	1.9 (0.8)	72.6 (27.0)	164.0 (90.2)
Trend*		T_{JT} = 356, z = 0.62, P = 0.535	T_{JT} = 295, z = −0.76, P = 0.450	T_{JT} = 356, z = 0.62, P = 0.535	T_{JT} = 363, z = 0.78, P = 0.436
$F_{(df)}$, P*		$F_{(4, 37)}$ = 1.92, P = 0.127	$F_{(4, 37)}$ = 0.98, P = 0.431	$F_{(4, 37)}$ = 4.66, P = 0.004	$F_{(4, 37)}$ = 2.06, P = 0.105

$F_{(df)}$, test statistic ANOVA; T_{JT}, test statistic Jonckheere-Terpstra test; t(df) = test statistic t test; z = standardized test statistic.
*Tested with <30 min collapsed with 30 to <60 min.

category (Table 4). For further consideration of the criterion validity for measuring sedentary behavior generally (not just accumulation), relationships of the seven self-report candidate measures with activPAL sitting %, standing %, and stepping % are shown in Supplemental Digital Content (see Table, Supplemental Digital Content 3, http://links. lww.com/JOM/A992, associations of self-report measures with activPAL workplace sitting, standing, and stepping). Six of the measures showed significant associations with sedentary behavior measured as activPAL % sitting, including typical day ($P < 0.001$) and conscious effort to reduce sitting ($P = 0.047$) that had not shown much criterion validity as measures of sitting accumulation.

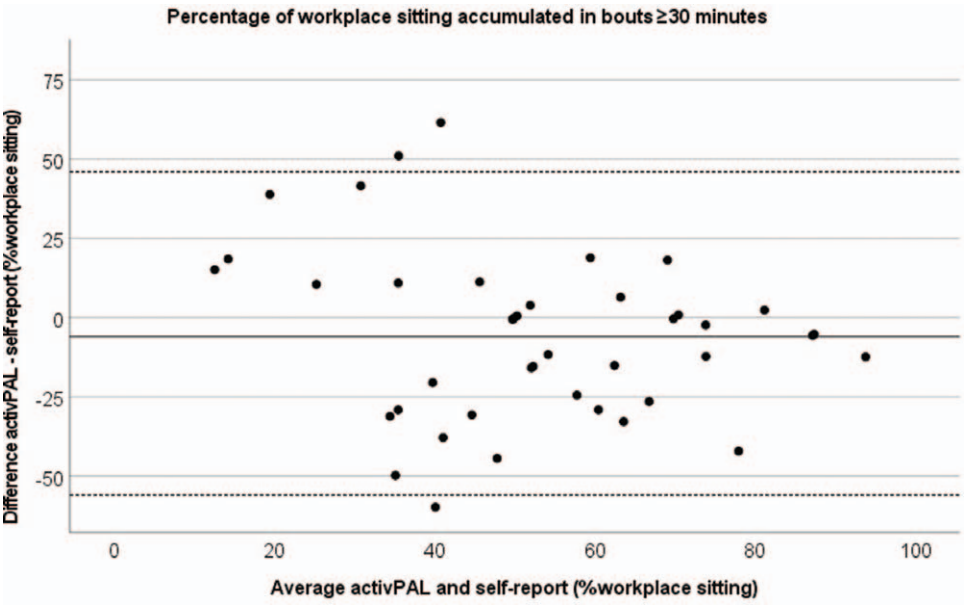

FIGURE 1. Bland Altman plot for percentage of workplace sitting accumulated in bouts more than or equal to 30 minutes as determined by self-report and activPAL. Solid line is mean difference and dashed lines are 95% limits of agreement. Mean difference (95% Limits of Agreement) = −5.8 (−57.7, 46.0).

TABLE 5. Summary of Findings for Survey Questions and Recommendations for Their Use in Measuring Sitting Time Accumulation

		Test-retest Findings		Validity Findings		Recommended Use			
Accumulation Question for Survey	**Measure Data Type**	**Reliability**	**Minimal Detectable Change**	**Correlation**	**Agreement**	**Cross-sectional**	**Cohort**	**Intervention**	**Comments**
% of sitting in periods ≥30 min	Continuous, 0–100	Y	Y	Y	Y	✓	✓	✓	Single item tailored to follow on from OSPAQ.
Interruption rate	Interval, 0–5+	N	N	Y	–	✓	≈✓*	≈✓*	Single item standalone measure. Multiple studies confirm correlation with a criterion.[17–19]
Ten item sitting strategy array									
Sitting strategy frequency	Interval, 0–100	Y	Y	Y	–	✓	✓	✓	Both measures are derived from the same array. Items could be added or removed based on knowledge of the study population.
Sitting strategy variety	Interval, 0–10	Y	Y	N	–	×≈✓†	×≈✓†	×≈✓†	
Longest sitting bout	Ordinal	N	–	N	–	×	×	×	
Conscious effort to reduce sitting	Binary	Y	–	N	–	≈✓†	≈✓†	≈✓†	Some validity for sedentary behavior more generally
Typical day	Ordinal or nominal	Y	–	N	–	≈✓‡	≈✓‡	≈✓‡	Also, Some validity for sedentary behavior more generally

Y adequate = significant finding ($P < 0.05$) or MDC% < 100; N not adequate; – not tested; ✓ potential for use; ≈✓ potential for use in selected circumstances; × recommend against use until validity is established.

*Suitable only when the sample size is large enough to offset the limited repeatability/large minimal detectable change.

†Unsuitable for measuring accumulation but suitable as a more general indicator of sedentary behavior based on relationship with workplace sitting time.

‡Suitable only with revised order of categories; also suitable as a more general indicator of sedentary behavior based on relationship with workplace sitting time.

DISCUSSION

This study was conducted to develop and assess measurement qualities of candidate self-report measures of accumulation of sitting time, specifically for those in work environments where work activities often involve sitting (office and call-center). The reliability and validity of the OSPAQ measures were similar to what has been seen in other studies[35,36] suggesting that the ability of this sample to recall their sedentary behavior is fairly typical of other working populations used in validity studies. A summary of the qualities of the measures of sitting accumulation and associated recommendations for their potential use is shown in Table 5. Among the seven candidate measures two brief and one longer measure were stand-out options in terms of performing well (compared with the other measures) in both reliability and validity for measuring sitting accumulation.

One of the candidate measures—percentage workplace sitting in long bouts—may be highly suited to studies (including interventions) that include the OSPAQ in their surveys and want a brief option. This single item measure was designed to immediately follow the OSPAQ questions about percentage time sitting, standing, and moving at work[21] and showed acceptable test-retest reliability, moderate correlation with the activPAL (ie, some ability to rank individuals), only a small average overestimation (ie, limited bias in group means) but with wide limits of agreement (ie, limited ability to capture the individual value identically to the activPAL). The revised single-item interruption rate question previously showed moderate correlations with device measured sit–stand or sedentary-movement transitions.[17–19] Here, the criterion validity was similar to previous studies, but reliability was worse[17–19] with correspondingly limited capacity to detect change. It may be a suitable choice for cross-sectional surveys that require a single, brief, stand-alone measure capable of ranking individuals or comparing groups as the relative degree to which they break up their sitting time. However, it would only be a good choice for interventions and cohort studies if the low repeatability is offset by a large sample size.

The 10 item array asking about strategies to break up sitting time offers a good choice for researchers to include in their questionnaires, if they are willing and able to include a lengthier option and are interested in obtaining more detailed contextual information about specific sitting behaviors that participants perform. Although both measures derived from this array showed good reliability (and a modest relative MDC), only sitting strategy frequency showed a significant, moderate correlation with the criterion, and is thus the preferred measure of sitting accumulation. Importantly, this approach captures accumulation indirectly and may be better termed a ''sitting breaks strategy score.'' Rather than assume the measure suits all workplaces, we would recommend researchers adapt their list of items to the strategies expected to be most common among their workers. In turn, we recommend researchers continue to collect and report on the strategies that workers are using to break up their sitting time, which serves an important function in guiding behavioral messages, even when studies can include more

accurate options, such as objective monitoring, to measure accumulation.

The three measures with binary and ordinal response format categorical responses had limited criterion validity as they did not show significant associations and trends respectively with the activPAL accumulation measures. While there was no trend in objectively measured accumulation for the typical day responses as asked in the survey, when the order of the categories was changed to reflect the expected amount of sitting there were significant trends in all accumulation outcomes. Asked in this way, the typical day measure may provide an indication of sitting time accumulation. While this study did not set out to identify measures of sedentary behavior more generally, based on their associations with workplace sitting time some of these measures may have utility. Workers who said they do not pay attention to their sitting time did show higher mean sitting time than those who reported they do pay attention and there was a trend in the expected direction for device-measured % sitting over categories of how workers usually spent their day. Both questions also showed moderate agreement between repeat administrations. There was no support for the longest sitting bout measure. Therefore, this question is not recommended for future research.

The strength of this study was the testing of several diverse candidate measures, allowing for comparison between the performances of these measures in the same sample. Another strength was the recruitment from multiple workplaces. However, the sample size was still relatively small and as such only capable of detecting statistically significant strong trends and large differences across categories. The sample size also affects the MDC (which reduces with increased sample size). The reporting of MDC was, however, a strength and this quality along with responsiveness to change is seldom available regarding sitting time questionnaires.[37] While this study design could assess only reliability, MDC, and criterion validity, ideally it would be good to follow up with intervention studies that can report on and validity for detecting changes and a responsiveness index that can be calculated with a two-group intervention design.[38] The MDC only reflects the degree of background variation in the measure (the noise), not the degree to which true changes are observed when they occur (the strength of the signal). A measure with a low MDC might not have good sensitivity to change if it is undermined by poor criterion validity.

The participants in this study were a mix of call center and office workers; therefore, they worked in typically high sitting environments.[9] However, the findings presented in this paper cannot be assumed to have external validity for all workplaces with high sitting. The external validity of questions regarding strategies to break up sitting in particular may depend on how similar the workplace environments are to the office settings whose workers' behaviors informed the list of strategies considered.[11,23] Additionally, the context for the validity study was workplaces that were taking part in a health and wellbeing program for which the workplaces volunteered. Measures were taken before intervention but after awareness raising about the program, so participants may have been particularly aware of their sitting behaviors, which may have inflated the findings. However, this did not seem to be the case since we did not see improvements in validity for the interruptions or OSPAQ measures compared with previous studies.[21,35,36] Future studies should assess the reliability, validity, and responsiveness to change of these measures of accumulation in varied workplace settings to ensure external validity of these findings.

CONCLUSION

This study identified several questions that might be suitable for researchers wanting to assess accumulation of workplace sitting time. The choice of appropriate question is dependent on the purpose for which it is intended, and guidance is provided within the paper.

ACKNOWLEDGMENTS

The researchers acknowledge the support of the BeUpstanding program and the time given by the participants in this study.

REFERENCES

1. Patterson R, McNamara E, Tainio M, et al. Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: a systematic review and dose response meta-analysis. *Eur J Epidemiol*. 2018;33:811–829.
2. Tremblay MS, Aubert S, Barnes JD, et al. Sedentary Behavior Research Network (SBRN) – terminology consensus project process and outcome. *Int J Behav Nutr Phys Activ*. 2017;14:75.
3. Saunders TJ, McIsaac T, Douillette K, et al. Sedentary behaviour and health in adults: an overview of systematic reviews. *Appl Physiol Nutr Metab*. 2020;45(10 (suppl)):S197–S217.
4. Katzmarzyk PT, Powell KE, Jakicic JM, Troiano RP, Piercy K, Tennant B. Sedentary behavior and health: update from the 2018 physical activity guidelines Advisory Committee. *Med Sci Sports Exerc*. 2019;51:1227–1241.
5. Zhai L, Zhang Y, Zhang D. Sedentary behaviour and the risk of depression: a meta-analysis. *Br J Sports Med*. 2015;49:705–709.
6. Chastin SFM, Egerton T, Leask C, Stamatakis E. Meta-analysis of the relationship between breaks in sedentary behavior and cardiometabolic health. *Obesity*. 2015;23:1800–1810.
7. Ross R, Chaput J-P, Giangregorio LM, et al. Canadian 24-hour movement Guidelines for adults aged 18–64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. *Appl Physiol Nutr Metab*. 2020;45(10 suppl 2):S57–S102.
8. Department of Health. Australia's Physical Activity and Sedentary Behaviour Guidelines and the Australian 24-Hour Movement Guidelines; 2019. Available at: https://www1.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines. Accessed May 6, 2021.
9. Toomingas A, Forsman M, Mathiassen SE, Heiden M, Nilsson T. Variation between seated and standing/walking postures among male and female call centre operators. *BMC Public Health*. 2012;12:154.
10. Parry S, Straker L. The contribution of office work to sedentary behaviour associated risk. *BMC Public Health*. 2013;13:296.
11. Ryde GC, Brown HE, Gilson ND, Brown WJ. Are we chained to our desks? Describing desk-based sitting using a novel measure of occupational sitting. *J Phys Act Health*. 2014;11:1318–1323.
12. Shrestha N, Kukkonen-Harjula KT, Verbeek JH, Ijaz S, Hermans V, Bhaumik S. Workplace interventions for reducing sitting at work. *Cochrane Database Syst Rev*. 2016;(3):CD010912.
13. Steeves JA, Bowles HR, McClain JJ, et al. Ability of thigh-worn ActiGraph and activPAL monitors to classify posture and motion. *Med Sci Sports Exerc*. 2015;47:952–959.
14. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc*. 2011;43:1561–1567.
15. Riviere F, Aubert S, Omorou AY, Ainsworth BE, Vuillemin A. Taxonomy-based content analysis of sedentary behavior questionnaires: a systematic review. *PLoS One*. 2018;13:e0193812.
16. Bakker EA, Hartman YAW, Hopman MTE, et al. Validity and reliability of subjective methods to assess sedentary behaviour in adults: a systematic review and meta-analysis. *Int J Behav Nutr Phys Activ*. 2020;17:75.
17. Pedisic Z, Bennie JA, Timperio AF, et al. Workplace Sitting Breaks Questionnaire (SITBRQ): an assessment of concurrent validity and test-retest reliability. *BMC Public Health*. 2014;14:1–9.
18. Sudholz B, Ridgers ND, Mussap A, Bennie J, Timperio A, Salmon J. Reliability and validity of self-reported sitting and breaks from sitting in the workplace. *J Sci Med Sport*. 2018;21:697–701.
19. Clark BK, Thorp AA, Winkler EA, et al. Validity of self-reported measures of workplace sitting time and breaks in sitting time. *Med Sci Sports Exerc*. 2011;43:1907–1912.
20. Healy GN, Goode AD, Abbott A, et al. Supporting workers to sit less and move more through the web-based beupstanding program: protocol for a single-arm, repeated measures implementation study. *JMIR Res Protoc*. 2020;9:e15756.
21. Chau JY, van der Ploeg HP, Dunn S, Kurko J, Bauman AE. Validity of the occupational sitting and physical activity questionnaire (OSPAQ). *Med Sci Sports Exerc*. 2012;44:118–125.
22. Katzmarzyk P, Church T, Craig C, Bouchard C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc*. 2009;41:998–1005.

23. Stephens SK, Eakin EG, Clark BK, et al. What strategies do desk-based workers choose to reduce sitting time and how well do they work? Findings from a cluster randomised controlled trial. *Int J Behav Nutr Phys Act*. 2018;15:98.
24. Sisson SB, Camhi SM, Church TS, et al. Leisure time sedentary behavior, occupational/domestic physical activity, and metabolic syndrome in U.S. men and women. *Metab Syndr Relat Disord*. 2009;7:529–536.
25. PAL Technologies. Why activPAL? 2021. Available at: https://www.palt.-com/why-activpal/. Accessed May 6, 2021.
26. Edwardson CL, Winkler EAH, Bodicoat DH, et al. Considerations when using the activPAL monitor in field-based research with adult populations. *J Sport Health Sci*. 2017;6:162–178.
27. Kringen NL, Healy GN, Winkler EAH, Clark BK. Accuracy of activPAL self-attachment methods. *Measure Phys Educ Exerc Sci*. 2016; 20:159–166.
28. Chastin SFM, Winkler EAH, Eakin EG, et al. Sensitivity to change of objectively-derived measures of sedentary behavior. *Measur Phys Educ Exerc Sci*. 2015;19:138–147.
29. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155–163.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
31. Polit DF. Getting serious about test-retest reliability: a critique of retest research and some recommendations. *Qual Life Res*. 2014;23:1713–1720.
32. Ries JD, Echternach JL, Nof L, Gagnon Blodgett M. Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Phys Ther*. 2009;89:569–579.
33. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19:231–240.
34. Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika*. 1954;41:133–145.
35. Jancey J, Tye M, McGann S, Blackford K, Lee AH. Application of the Occupational Sitting and Physical Activity Questionnaire (OSPAQ) to office based workers. *BMC Public Health*. 2014;14:762.
36. van Nassau F, Chau JY, Lakerveld J, Bauman AE, van der Ploeg HP. Validity and responsiveness of four measures of occupational sitting and standing. *Int J Behav Nutr Phys Act*. 2015;12:1–9.
37. Dall P, Coulter E, Fitzsimons C, et al. TAxonomy of Self-reported Sedentary behaviour Tools (TASST) framework for development, comparison and evaluation of self-report tools: content analysis and systematic review. *BMJ Open*. 2017;7:e013844.
38. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*. 1987;40:171–178.

