Data Documentation

November 16, 2021

1 Participants and days excluded from all analyses

In this documentation, we refer to the date when the participant was scheduled to have their second in-person lab visit (i.e., their Quit Day) as 'Day 0' and subsequent days as 'Day 1', 'Day 2', ..., and so on. Day 0 was the first day when participants could be micro-randomized. Day 0 marks a significant transition in the *Sense2Stop* study: only participants who attended their second in-person lab visit were able to have study coordinators activate the *mCerebrum* app's micro-randomization capabilities; *mCerebrum* is an app installed on the mobile phones loaned to participants and orchestrates data collection between the mobile phone and wearables. The last day when participants could be micro-randomized occurred 10 days after Day 0 (i.e., on 'Day 10'). Thus, each participant would have at most 11 days when they could have been micro-randomized. 'Day 10' was also the day prior to the participants' third lab visit, the date when participants were asked to return the mobile phone and wearables loaned to them.

Table 1: Participants and days excluded from all analysis

No. Participants	No. Days/Participant	Total Participant-Days	Rationale
5	11	55	Pilot Participants
12	11	132	Did not attend their second lab visit
1	6	6	Withdrew from the study on Day 6

The remainder of this document exclude the participants and days described above.

2 Availability for micro-randomization

Unavailability due to inclement weather

Although participants generally attended their second in-person lab visit on their scheduled date, there were few instances when participants contacted study coordinators to delay their second in-person lab visit. In all these cases, participants reported that they could not attend their second in-person lab visit on Day 4 due to hazardous weather conditions, but were able to complete their second in-person lab visit on Day 5. We regarded all decision points on Day 0 as unavailable for these participants. *Rationale:* Study protocol (Section 9.2, Version #24, 2018-01-12) dictates that during their first in-person lab visit, participants will plan to taper their smoking on the day just prior to their scheduled second in-person clinic visit; participants are asked in advance to plan to engage only in single puff smoking episodes on that day. Hence, regardless of when participants actually completed their second in-person lab visit, we expect participant's smoking behavior on the day of their scheduled second in-person lab visit to reflect a legitimate attempt to quit smoking. Hence, we do not set the *actual* date of their second in-person visit as Day 0; rather, we set the date of their *scheduled* second in-person clinic visit as Day 0. In this way, Day 0 can be interpreted as 'Quit Day' for all participants in the study.

Unavailability due to omission of initiating start-of-day

Micro-randomizations for a particular day of the study (e.g., on Day 9) did not occur unless the participant initiated 'start-of-day' through pressing a start-of-day icon within the *mCerebrum* app. In other words, for micro-randomizations to occur daily between Day 0 to Day 11, participants needed to press the start-of-day icon within the *mCerebrum* app daily after waking. We regarded all decision points on those participant-days when the participant did not press the start-of-day icon within the *mCerebrum* app as unavailable.

Table 2: Participant-days when no decision points were regarded as available for micro-randomization

No. Days/Participant	No. Participants	Total Participant-Days
Inclement weather		
1	2	2
Start-of-day was not initiated		
10	2	20
9	4	36
8	4	32
7	2	14
6	2	12
5	1	5
4	3	12
3	1	3
2	5	10
1	8	8

Unavailability for micro-randomizations determined in real-time by the mCerebrum app

Only the 720 minutes following initiation of start-of-day can be potentially regarded by the *mCerebrum* app as available for micro-randomization. We refer to start-of-day as DP 0, and the 1, 2, . . . , 720 minutes after start-of-day as DP 1, 2, 720. Hence, the upper limit in the number of available decision points (DP's) for each participant-day is 721.

Availability for micro-randomization is determined in real-time by the mCerebrum app via a decision rule (see Figure 2 in Battalio, et al., 2021) involving consideration of the following conditions as long as a particular DP (among DP 0, DP 1, DP 720) does not fall after 'sleep time'.

In brief, 'sleep time' refers to the time of day at which the *mCerebrum* app automatically paused Random EMA and EMI delivery until the next time start-of-day is initiated; 'sleep time' was pre-set by study staff in consultation with the participant prior to completion of the first in-person clinic visit. Hence, DP's which occur after 'sleep time', but prior to the next time start-of-day was initiated will be regarded as unavailable for micro-randomization.

- Whether the participant was at the *peak* of an episode
- Whether less than 50% of minutes between the *start* and *peak* of an episode could be regarded as having poor data quality
- Whether the episode was classified as *stress* or *not stress*
- Whether privacy mode was not activated
- Whether time since last Random EMA was more than 10 minutes
- Whether time since last participant-initiated EMA was more than 10 minutes
- Whether the participant was not physically active in the last 5 minutes
- Whether the participant was not driving in the last 5 minutes
- Whether the mobile phone's battery was greater than 10% in the last 5 minutes

Table 3: Among participant-days not included in Table 1 or Table 2, count decision points (DP's) regarded by the mCerebrum app as available for micro-randomization.

No. DP's/Participant-Day	No. Participant-Days	Total DP's
48	1	48
47	1	47
45	1	45
43	2	86
41	2	82
40	1	40
38	2	76
37	2	74
36	3	108
35	3	105
34	1	34
33	2	66
32	2	64
31	5	155
30	7	210
29	3	87
28	9	252
27	12	324
26	6	156
25	7	175
24	10	240
23	6	138
22	13	286
21	8	168
20	8	160
19	13	247
18	18	324
17	13	221
16	10	160
15	9	135
14	10	140
13	13	169
12	8	96
11	16	176
10	11	110
9	10	90
8	13	104
7	8	56
6	10	60
5	10	50
4	18	72
3	6	18
2	10	20
1	15	15
0	140	0
Grand Total:		
		5489

Table 4: Among participant-days not included in Table 1 or Table 2, count decision points (DP's) regarded by the mCerebrum app as available for micro-randomization. In contrast with Table 3, the current table does not count the last 720 - M DP's; here, M=120 minutes. In other words, the last 720 - 120 = 600 DP's will not be included in the analytic dataset, i.e., these DP's were excluded.

No. DP's/Participant-Day	No. Participant-Days	Total DP's
41	1	41
40	1	40
39	1	39
36	2	72
33	3	99
32	5	160
31	1	31
30	2	60
29	5	145
28	3	84
27	7	189
26	4	104
25	5	125
24	7	168
23	4	92
22	11	242
21	13	273
20	13	260
19	13	247
18	15	270
17	20	340
16	9	144
15	9	135
14	13	182
13	13	169
12	9	108
11	18	198
10	18	180
9	12	108
8	12	96
7	9	63
6	13	78
5	14	70
4	19	76
3	8	24
2	9	18
1	14	14
0	143	0
Grand Total:		
		4744

3 Working with predictions of a stress episode detection algorithm

Decision rules to link stress/not stress/physically active episodes to specific participant-days

In this section, given a specific participant-day, let t_0 denote the time-of-day when the participant initiates start-of-day and $t_0 + 720$ be 720 minutes following the moment when the participant initiates start-of-day. Stress/not stress/physically active episodes are linked to participant-days in which start-of-day was initiated, i.e., we do not attempt a linkage of episodes to participant-days in which start-of-day was not initiated.

Among participant-days not included in Table 1 or Table 2, a stress/not stress/physically active episode was linked to a specific participant-day if the episode:

- **(Case 1)** Began and ended within time time interval $[t_0, t_0 + 720]$, i.e., $A, C \in [t_0, t_0 + 720]$
- **(Case 2)** Began after start-of-day was initiated, i.e., $A \in [t_0, t_0 + 720]$, but ended after 720 minutes following the moment when the participant initiates start-of-day, i.e., $C \notin [t_0, t_0 + 720]$
- **(Case 3)** Began prior to when start-of-day was initiated, i.e., $A \notin [t_0, t_0 + 720]$, but ended prior to 720 minutes following the moment when the participant initiates start-of-day, i.e., $C \in [t_0, t_0 + 720]$

Table 5: No. Episodes Classified as Stressed/Not Stressed/Physically Active

Episode Classification	No. Episodes
Stress	1137
Physically Active	4642
Not Stress	11251

We note that the above approach of linking episodes to participant-days generally results in each episode being linked to one and only one participant-day. Exceptions are 19 episodes which are each linked to 2 participant-days. These episodes belong to either Case 2 or Case 3 above and range between 5.27 and 40.03 hours long, i.e., hours elapsed between A and C range between 5.27 and 40.03 hours.

Linking episode-level stress/not stress/physically active classifications to minute-by-minute classifications

Each minute between t_0 and $t_0 + 720$ is labelled as a stress/not stress/physically active minute if it lies within a stress/not stress/physically active episode. Episodes considered are only those which have been linked to the specific participant-day using the rules described above.
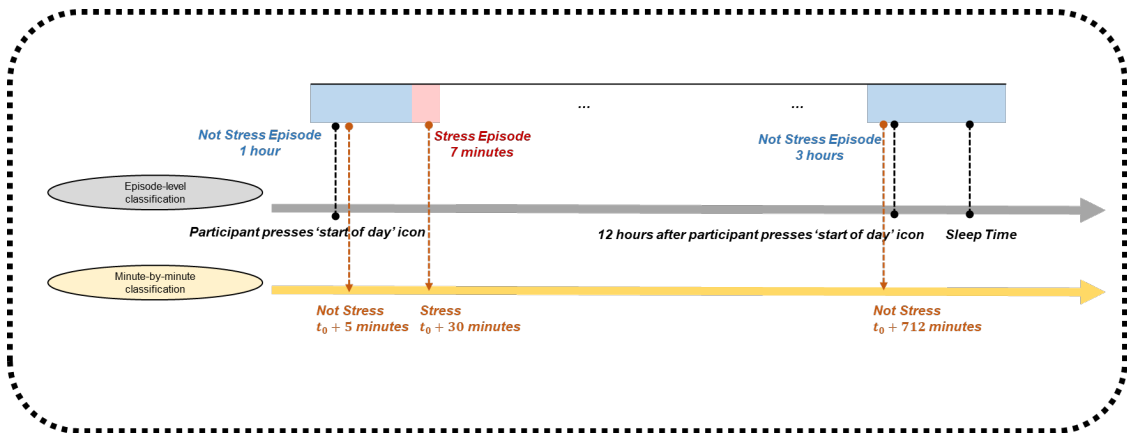

Figure 1: A visual depiction of how episode-level stress/not stress/physically active classifications are linked to minute-by-minute stress/not stress/physically active classifications in one hypothetical participant-day.

How often and how long were the Stress/Not Stress/Physically Active Episodes?

Table 6: Minutes elapsed between the start, peak, and end of stress episodes, summarized in percentiles

	0% (MIN)	10%	50% (MEDIAN)	90%	100% (MAX)
Start to Peak (A to B)	0.00	1.00	3.99	7.42	4115.01
Peak to End (B to C)	0.58	2.99	6.01	11.99	841.01
Start to End (A to C)	1.82	5.00	10.00	17.99	4116.01

Table 7: Minutes elapsed between the start, peak, and end of not stress episodes, summarized in percentiles

	0% (MIN)	10%	50% (MEDIAN)	90%	100% (MAX)
Start to Peak (A to B)	0.00	1.01	4.00	12.98	809.00
Peak to End (B to C)	0.66	1.00	4.01	9.13	985.99
Start to End (A to C)	0.91	4.00	9.01	19.01	986.99

4 Overlapping windows?

We use DP's included in Table 4 as our starting point. In contrast to those DP's in Table 3, we excluded DP's $\{720 - m : m = 1, 2, 3, ..., M\}$, where $M = 120$. In other words, we additionally excluded the DP's for which we will not be able to 'look ahead' for a full $M = 120$ minutes.

Table 8: Among DP's regarded as available for micro-randomization in Table 4: How many were followed by any microrandomization within the next M=120 minutes?

At current DP (k)		Among current DP's (k), in how many did any of the following occur within the next M=120 minutes (i.e., (k+1, k+1, ..., k+M))?	
Randomization assignment	Total Number of DP's	Without any micro-randomization	With any micro-randomization
Micro-randomized to 'No Prompt'	4160	211	3949
Micro-randomized to 'Prompt'	584	131	453
Grand Total:			
	4744	342	4402

Hence, Table 8 shows that among the 4744 DP's regarded as available for micro-randomization in the kth DP, 4402 were followed by a micro-randomization within the next $M = 120$ minutes. This is further broken down:

- Among the 4160 DP's micro-randomized to 'No Prompt' in the kth DP, 3949 were followed by a micro-randomization within the next $M = 120$ minutes.
- Among the 584 DP's micro-randomized to 'Prompt' in the kth DP, 453 were followed by a micro-randomization within the next $M = 120$ minutes.

Next, we break down the last column of Table 8.

Table 9: Among DP's regarded as available in Table 4 and were followed by any micro-randomization within the next M=120 minutes, how many were micro-randomized to 'Prompt' within the next M=120 minutes?

At current DP (k)		Among current DP's (k), in how many did any of the following occur within the next M=120 minutes (i.e., (k+1, k+1, ..., k+M))?	
Randomization assignment	Total Number of DP's	Never Micro-randomized to 'Prompt'	Micro-randomized to 'Prompt'
Micro-randomized to 'No Prompt'	3949	2333	1616
Micro-randomized to 'Prompt'	453	325	128
Grand Total:			
	4402	2658	1744

Hence, Table 9 shows that among the 4402 DP's regarded as available for micro-randomization in the kth DP and were followed by a micro-randomization within the next $M = 120$ minutes, 1744 were micro-randomized to 'Prompt' within the next $M = 120$ minutes. This is further broken down:

- Among the 3949 DP's micro-randomized to 'No Prompt' in the kth DP, 1616 were followed by a micro-randomization to 'Prompt' within the next $M = 120$ minutes.
- Among the 453 DP's micro-randomized to 'Prompt' in the kth DP, 128 were followed by a micro-randomization to 'Prompt' within the next $M = 120$ minutes.

5   References

1. Battalio, S. L., Conroy, D. E., Dempsey, W., Liao, P., Menictas, M., Murphy, S., ... & Spring, B. (2021). Sense2Stop: A micro-randomized trial using wearable sensors to optimize a just-in-time-adaptive stress management intervention for smoking relapse prevention. Contemporary Clinical Trials, 109, 106534.