

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## **A digital epidemiological and citizen science methodology to capture prospective physical activity within social and physical contexts: A SMART platform study**

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-036787
Article Type:	Original research
Date Submitted by the Author:	18-Jan-2020
Complete List of Authors:	Katapally , Tarun ; University of Regina, ; University of Saskatchewan, Chu, Luan; University of Saskatchewan,
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, SPORTS MEDICINE

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3 1 **Title:** A digital epidemiological and citizen science methodology to capture prospective physical  
4 2 activity within social and physical contexts: A SMART platform study

5 3  
6 4 Authors: Tarun Reddy Katapally, MS, PhD,<sup>1,2,3</sup>, Luan Manh Chu, MS<sup>3,4</sup>  
7 5  
8  
9

10 6 <sup>1</sup>Johnson Shoyama Graduate School of Public Policy, University of Regina, 2155 College Ave,  
11 7 Regina, SK S4M0A1, Canada. Tarun.katapally@uregina.ca.  
12 8

13 9 <sup>2</sup>Johnson Shoyama Graduate School of Public Policy, University of Saskatchewan, 101  
14 10 Diefenbaker Pl, Saskatoon, SK S7N 5B8, Canada. Tarun.katapally@usask.ca.  
15 11

16 12 <sup>3</sup>College of Medicine, Health Science Building, 107 Wiggins Road, University of Saskatchewan,  
17 13 Saskatoon, SK S7N 5E5, Canada. Tarun.katapally@usask.ca.  
18 14

19 15 <sup>4</sup>Canadian Centre for Health and Safety in Agriculture, University of Saskatchewan, 104 Clinic  
20 16 Place, PO Box 23, Saskatoon, SK S7N 2Z4, Canada. Cml779@mail.usask.ca  
21 17  
22 18  
23 19  
24 20  
25 21  
26 22  
27 23  
28 24  
29 25  
30 26  
31 27  
32 28  
33 29  
34 30  
35 31  
36 32  
37 33  
38 34  
39 35  
40 36

41 27 Corresponding author: Tarun Reddy Katapally  
42 28 Johnson Shoyama Graduate School of Public Policy, University of Regina  
43 29 110-2 Research Drive, Innovation Place, Regina, SK, S4S7H9, Canada  
44 30 Phone: 1 3065854544  
45 31 Email: [Tarun.katapally@uregina.ca](mailto:Tarun.katapally@uregina.ca)  
46 32  
47 33  
48 34  
49 35  
50 36  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 37 **ABSTRACT**  
4

5 38 **Objectives** The purpose of this study was to develop a novel and replicable methodology of  
6  
7 39 mobile ecological momentary assessments (EMAs) to capture prospective physical activity (PA)  
8  
9 40 within free-living social and physical contexts by leveraging citizen-owned smartphones running  
10  
11 41 on both Android and iOS systems.  
12  
13

14 42 **Design** Data were obtained from the cross-sectional pilots of the SMART Platform, an  
15  
16 43 innovative citizen science and mobile health initiative for active living surveillance.  
17  
18

19 44 **Setting** The study was conducted in the cities of Regina and Saskatoon, Canada.  
20  
21

22 45 **Participants** 538 citizen scientists ( $\geq 18$  years) provided PA data during 8 consecutive days using  
23  
24 46 a custom-built smartphone app. Citizen scientists who completed daily time-triggered EMAs  
25  
26 47 (capturing prospective PA) and International Physical Activity Questionnaire (IPAQ) were  
27  
28 48 included in the final analyses.  
29  
30

31 49 **Outcome measures** EMAs enabled reporting of light, moderate, and vigorous PA, as well as  
32  
33 50 physical and social contexts of PA via complex looped linking of intensity and context questions.  
34  
35 51 Retrospective PA was reported using IPAQ. For both measures, PA intensities were categorized  
36  
37 52 into mean light and moderate-to-vigorous PA/day. Wilcoxon signed ranks tests and Spearman  
38  
39 53 correlation procedures were conducted to compare PA intensities reported via EMAs and IPAQ.  
40  
41

42 54 **Results** The findings showed discrepancies between EMA and IPAQ measures of PA. Daily  
43  
44 55 time-triggered EMAs were able to capture not only prospective light and moderate-to-vigorous  
45  
46 56 PA, but also enabled PA reporting across varied physical and social contexts. Among physical  
47  
48 57 contexts, citizen scientists reported accumulating PA predominantly at home. Among social  
49  
50 58 contexts, citizen scientists reported accumulating PA predominantly by themselves.  
51  
52  
53

1  
2  
3 59 **Conclusions** These findings suggest that time-triggered mobile EMAs are an effective method to  
4  
5 60 record comprehensive prospective PA accumulated across multiple physical and social contexts.  
6  
7 61 With over 3 billion smartphones users globally, these ubiquitous tools can be leveraged via  
8  
9 62 citizen science to understand active living patterns of large populations in free-living conditions  
10  
11 63 using EMAs.  
12  
13

14 64 **Keywords:** Physical activity, mHealth, Ecological Momentary Assessments, Measurement,  
15  
16 65 Citizen Science, Digital Epidemiology  
17  
18  
19

### 20 66 **Strengths**

- 21 67 • This study addresses current discrepancies in mobile ecological momentary assessment  
22  
23 68 (EMA) methodologies (e.g., triggering processes, time to follow-up), as well as  
24  
25 69 limitations in terms of usage of identical mobile devices need to be addressed to deploy  
26  
27 70 EMA among large populations.  
28  
29  
30 71 • This study shows that time-triggered mobile EMAs are an effective method to record  
31  
32 72 comprehensive daily prospective physical activity.  
33  
34  
35 73 • This study shows that EMAs can be used to capture both physical and social context of  
36  
37 74 physical activity prospectively.  
38  
39

### 40 75 **Limitations**

- 41  
42 76 • The main limitation is the small sample size after applying the inclusion criteria  
43  
44  
45  
46 77

## 48 78 **INTRODUCTION**

49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 79 Advances in mobile technology over the past decade have facilitated the innovation of  
4  
5 80 ecological momentary assessments (EMAs), which are digital epidemiological tools that aid in  
6  
7 81 understanding environmental, social, and behavioural processes.<sup>1,2</sup> EMAs can capture real-time  
8  
9 82 data that reflect the dynamics of participants' experiences in their natural environment and thus  
10  
11 83 they are increasingly being used to monitor health behaviors among populations across the life  
12  
13 84 course.<sup>3-5</sup> In active living research, evidence indicates that EMAs are a valid, reliable, and  
14  
15 85 feasible method of data collection.<sup>6,7</sup>  
16  
17  
18  
19

20 86 EMAs are an advancement over traditional self-report methods as they enable data collection  
21  
22 87 more proximal to the time and place that a behavior has occurred.<sup>2,8</sup> Moreover, EMAs overcome  
23  
24 88 many of the limitations of traditional self-report surveys to provide information regarding  
25  
26 89 specific activity types (e.g. watching TV vs video gaming) and capture important factors that  
27  
28 90 influence health behaviors such as mood and environmental perceptions.<sup>5,9,10</sup> In measuring  
29  
30 91 physical activity (PA) intensities, EMAs have been shown to minimize recall<sup>6,11</sup> and social  
31  
32 92 desirability bias<sup>12</sup> of traditional self-report measures.  
33  
34  
35  
36

37 93 Several studies have examined the validity of smartphone-based EMAs compared to other  
38  
39 94 objective devices (accelerometers, pedometers) and self-report measures of PA.<sup>2,10-16</sup> Overall,  
40  
41 95 estimates from EMAs were found to be highly correlated with accelerometer estimates.<sup>13,14</sup>  
42  
43 96 However, this evidence also indicates that PA was over-reported when International Physical  
44  
45 97 Activity Questionnaire (IPAQ) was used<sup>2</sup> and that daily PA EMA reports were not significantly  
46  
47 98 associated with their traditional recall measures.<sup>12</sup>  
48  
49  
50  
51  
52  
53

1  
2  
3 99 Currently there is little evidence of existing EMA methods that capture PA intensities across  
4  
5 100 various physical (leisure-time PA, transit-related PA, occupation-related PA, and  
6  
7 101 household/domestic-related PA)<sup>18</sup> and social contexts (with family, friends etc.).<sup>15</sup> Moreover,  
8  
9 102 there are discrepancies in smartphone-based EMA methodologies, which range from inconsistent  
10  
11 103 EMA triggering processes and varying times of prospective follow-up, to limitations of using  
12  
13 104 identical mobile devices and operating systems.<sup>19</sup>  
14  
15  
16  
17

18 105 The objective of this study is to address current deficiencies in PA EMA approaches by  
19  
20 106 developing a novel and replicable citizen science methodology of standardized time-triggered  
21  
22 107 smartphone-based EMAs to capture prospective PA within free-living social and physical  
23  
24 108 contexts by leveraging citizen-owned smartphones running on both Android and iOS systems.  
25  
26 109 This study will also compare EMA measures with traditional self-report measures of PA within  
27  
28 110 the same cohort.  
29  
30  
31

## 32 111 **METHODS**

### 33 34 35 112 **Design**

36  
37  
38 113 This study is part of the SMART Platform, which is a mobile health (mHealth) and citizen  
39  
40 114 science initiative for active living surveillance, integrated knowledge translation, and policy and  
41  
42 115 real-time interventions.<sup>8,16,17</sup> Citizen science is a participatory approach where participants,  
43  
44 116 termed citizen scientists, actively engage in the research process from data collection to  
45  
46 117 knowledge translation, thus improving the probability of longitudinal participant compliance.<sup>18</sup> A  
47  
48 118 detailed description of SMART Platform's methods, including recruitment and data collection  
49  
50 119 strategies, are described in the Platform's methodology publication.<sup>19</sup>  
51  
52  
53



1  
2  
3 120 The data for this study have been obtained from the 2017 (April 1 to May 31) and 2018  
4  
5 121 (January 4 to March 31) cohorts of the SMART Platform,<sup>8</sup> which is a prospective investigation  
6  
7 122 designed to capture active living data from adults residing in the two largest urban centers in  
8  
9 123 Saskatchewan, Canada (Regina and Saskatoon). All subjective (via traditional validated surveys  
10  
11 124 and EMAs) and objective data (via smartphones sensors) related to PA, sedentary behaviour, and  
12  
13 125 perception of environment, individual motivation, health outcomes, and eudaimonic well-being  
14  
15 126 were obtained through citizen-owned smartphones on 8 consecutive days (**Figure 1**).

### 17 127 **Patient and public involvement**

18  
19  
20 128 Participants in the SMART Platform are “citizen scientists” as they can engage with the  
21  
22 129 researchers at all stages of the research process. Thus, citizen scientists informed the design,  
23  
24 130 research questions and outcome measures. As part of the social media campaign for recruitment,  
25  
26 131 citizen scientists were encouraged to inform their friends about the study. Finally, as integrated  
27  
28 132 knowledge translation is part of the SMART Platform, results are disseminated throughout the  
29  
30 133 study period using the community voices webpage of the Platform’s website:  
31  
32 134 <https://www.smartstudysask.com/community-voices>

### 33 135 **Recruitment and participants**

34  
35  
36 136 Citizen scientists for SMART Adult cohorts were recruited online through social media, and  
37  
38 137 in-person from the universities of Regina and Saskatchewan and community centres located in  
39  
40 138 different neighbourhoods in each city to capture a socioeconomically representative sample.  
41  
42 139 Citizen scientists were guided to download Ethica (Ethica Data Services Inc.), an  
43  
44 140 epidemiological smartphone application (app), specifically adapted for the SMART Platform,  
45  
46 141 which captures data through both Android and iOS platforms. All citizen scientists provided

1  
2  
3 142 informed consent through the app and confirmed their age ( $\geq 18$  years) before joining the study.  
4  
5 143 Ethics approval was obtained from the universities of Regina and Saskatchewan through a  
6  
7  
8 144 synchronized review protocol (REB # 2017-29).  
9

## 10 145 **Measures**

11  
12  
13 146 The two primary measures used in this study are the International Physical Activity  
14  
15 147 Questionnaire (IPAQ),<sup>20</sup> which collects retrospective PA in 4 physical domains (recreation,  
16  
17 148 active transportation, work, and home), and the SMART Platform's modified EMA, which  
18  
19  
20 149 captures prospective daily PA in both social and physical contexts.  
21  
22

### 23 150 *IPAQ*

24  
25  
26 151 IPAQ was deployed at baseline as soon as citizen scientists downloaded the app to self-report  
27  
28 152 physical activities over the past 7 days that were of at least 10 consecutive minutes in duration.  
29  
30 153 These activities were categorized by 4 domains: 1) Recreation (e.g., weight training, sports  
31  
32 154 (soccer, hockey, etc.), aerobics, running, jogging, swimming, cycling, etc.); 2) Household (e.g.,  
33  
34 155 carrying light loads, sweeping, washing windows, and raking, etc.); 3) Transportation (e.g.,  
35  
36 156 travelling in a train, bus, car, or other kind of motor vehicle, etc.); and 4) Work (e.g. heavy  
37  
38 157 lifting, digging, heavy construction or climbing upstairs, etc.). The records included the number  
39  
40 158 of times per week (within the last 7 days) and average minutes per day for each activity.  
41  
42  
43  
44  
45 159  
46  
47  
48 160  
49  
50

### 51 161 *Adapted Daily EMAs*

1  
2  
3 162 Using the SMART platform, time-triggered modified EMAs (Figure 1) were developed,  
4  
5 163 tested, and piloted, before being pushed to citizen scientists' smartphones between 8pm and  
6  
7 164 8:30pm on each day for 8 consecutive days. These EMAs were designed to expire at 3am the  
8  
9 165 next day. Citizen scientists were asked to report only those physical activities that were of at  
10  
11 166 least 10 minutes in duration at a time. More importantly, each EMA was designed to not only  
12  
13 167 measure intensity and volume (in minutes) of PA, but also to capture social (i.e., with whom they  
14  
15 168 accumulated PA [Figure 1C]) and physical contexts (i.e., where they accumulated PA [Figure  
16  
17 169 1D]). This design was achieved by creating a looped linkage, where upon entering the volume of  
18  
19 170 each activity, the EMA triggered the social and physical context questions.  
20  
21  
22  
23

#### 24 171 **Derived variables – Intensities and Volume of PA**

##### 25 172 *IPAQ*

26  
27  
28 173 Thirty-seven questions related to PA were asked and 3 different categories of intensities were  
29  
30 174 created (light, moderate, and vigorous PA) by combining PA across 4 domains: recreation,  
31  
32 175 household, workplace, and active transportation. Moderate and vigorous PA intensities are  
33  
34 176 combined to derive “moderate-to-vigorous PA.” After conducting several aggregation  
35  
36 177 techniques, 2 final intensity variables were derived for IPAQ retrospective PA: mean minutes per  
37  
38 178 day of light and moderate-to-vigorous PA.  
39  
40  
41  
42  
43  
44

45 179

46 180

##### 47 181 *Adapted Daily EMAs*

1  
2  
3 182 A similar approach was employed to derive two final intensity variables for EMA prospective  
4  
5 183 PA: mean minutes per day of light and moderate-to-vigorous PA. For example, the light PA  
6  
7 184 included walking, light hiking, any light physical activity/sport (e.g. golf bowling etc.), yoga, and  
8  
9 185 light intensity household chores (e.g. washing dishes sweeping laundry gardening). Moderate-to-  
10  
11 186 vigorous PA included moderate to vigorous hiking, running, biking, any team sport (football  
12  
13 187 hockey soccer etc.), any other sport or activity (swimming canoeing skiing etc.), weight training,  
14  
15 188 dance/aerobic/cardio exercise, and moderate-to-vigorous intensity household chores (e.g.  
16  
17 189 shovelling driveways, washing a car etc.).  
18  
19  
20  
21

### 22 190 *Physical Context*

23  
24  
25 191 PA information from the IPAQ and EMAs (based on the question "Where did you do this  
26  
27 192 activity?") were grouped into domains. Domain 1: PA at workplace (IPAQ) and from work  
28  
29 193 (EMAs). Domain 2: Transportation PA (IPAQ) and from street (EMAs). Domain 3: Housework,  
30  
31 194 house maintenance, and caring from family (IPAQ) and from home (EMAs). Domain 4:  
32  
33 195 Recreation, sport, and leisure-time PA (IPAQ) and from park, gym, and sport facility (EMAs).  
34  
35  
36

### 37 196 *Social Context*

38  
39  
40 197 Social context information was collected via EMA question, "With whom did you do this  
41  
42 198 activity?" for each physical activity that the participants reported. Categories for social context  
43  
44 199 included "by myself, with my dog, with my friend(s), with my parent(s)," among others.  
45  
46  
47

48 200

### 51 201 **Statistical analyses**

202 The inclusion criterion to determine the final sample was dependent on citizen scientists  
203 completing the IPAQ, and answering the daily EMA on at least 3 days. Continuous estimates  
204 were reported as means with standard deviations (SD) and medians with 25<sup>th</sup> and 75<sup>th</sup> percentiles,  
205 depending on normality. Where estimates were non-normal and positively skewed, median and  
206 interquartile ranges were used. Wilcoxon signed ranks tests and Spearman correlation procedures  
207 were conducted to compare PA intensities and domain-based PA reported via IPAQ and EMAs.  
208 Correlation coefficient values of <0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80 and 0.81-1.0 were  
209 considered as weak, fair, moderate, strong and very strong correlation, respectively.<sup>21</sup> Analyses  
210 were conducted in SPSS version 24.0 (SPSS Inc., Chicago IL, USA) with significance set at  
211 alpha < 0.05.

## 212 RESULTS

213 After applying the decision rule of including only those citizen scientists who completed  
214 IPAQ, and answered the daily EMA on at least 3 days, out of 538 participants, only 89 were  
215 included in this study, among whom 47 identified as female (52.80%). The final sample had the  
216 mean age of 36.7 years (SD=15.74), and a mean body mass index of 28.34 (SD=7.82). The  
217 median (25<sup>th</sup>, 75<sup>th</sup> percentiles) and the mean (SD) duration of time (minutes per day) spent in  
218 each of the activity intensities (light, moderate and vigorous), as well as overall PA were derived  
219 from both IPAQ and EMA measures.

220 Using EMAs, citizen scientists reported 140.91, 87.16, and 70.38 mean minutes/day of  
221 overall PA, light PA, and moderate-to-vigorous PA. The same citizen scientists reported 194.39,  
222 116.99, and 98.42 mean minutes/day of overall PA, light PA, and moderate-to-vigorous PA  
223 using the IPAQ (**Tables 1 and 2**). These findings show that although there are no significant

1  
2  
3 224 differences between activity intensities reported via EMAs and IPAQ, citizen scientists  
4  
5 225 consistently overestimated their PA using IPAQ. **Table 3** demonstrates the correlation between  
6  
7 226 EMA and IPAQ measures to show that overall PA ( $\rho=0.414$ ,  $p<0.001$ ), and light ( $\rho =0.261$ ,  
8  
9  $p=0.012$ ) and moderate-to-vigorous PA ( $\rho =0.316$ ,  $p=0.009$ ) were fairly correlated. **Figure 2**  
10  
11 227 shows the visual representation of these correlations.  
12  
13 228

14  
15 229 **Figure 3 and 4** demonstrate the distribution of overall PA accumulated across different  
16  
17 230 physical and social contexts, as reported by citizen scientists using EMAs. Among physical  
18  
19 231 contexts, citizen scientists reported accumulating overall PA predominantly at home (26.4%), on  
20  
21 232 the streets ([20.4%] i.e., active transportation), at the gym (13.7%), at work (13.1%), and in parks  
22  
23 233 (12.3%). When it comes to social context, citizen scientists overwhelmingly reported  
24  
25 234 accumulating overall PA by themselves (64.2%), with some reporting being active with friends  
26  
27 235 (14.7%) and relatives (6.3%).  
28  
29

30  
31 236 As IPAQ captures PA in 4 physical domains (workplace, active transportation, household,  
32  
33 237 and recreation, sport and leisure-time) to compare estimates between EMA and IPAQ, EMA  
34  
35 238 estimates of overall PA accumulated across various physical contexts were categorized to match  
36  
37 239 the physical domains of IPAQ. Using EMAs, citizen scientists reported 20.50, 16.41, 25.33, and  
38  
39 240 20.88 mean minutes/day of overall PA across workplace, active transportation, household, and  
40  
41 241 recreation, sport and leisure-time domains, respectively. Using IPAQ, the same citizen scientists  
42  
43 242 reported 32.14, 43.97, 38.27, and 145.90 mean minutes/day of overall PA across workplace,  
44  
45 243 active transportation, household, and recreation, sport and leisure-time domains, respectively.  
46  
47  
48  
49

50 244 These findings show that there is a consistent pattern of over-reporting of overall PA across  
51  
52 245 all physical domains when citizen scientists used IPAQ, with statistically significant differences  
53

246 observed in active transportation ( $p=0.002$ ) and recreation, sport and leisure-time domains  
247 ( $p=0.003$ ) (**Table 4**). Corroborating these findings, **Table 5** demonstrates correlation between  
248 EMA and IPAQ physical domain measures, with moderate correlation being depicted with  
249 overall PA accumulated in household domain light ( $p=0.607$ ,  $p=0.036$ ).

## 250 **DISCUSSION**

251 The objective of this study was to address current deficiencies in PA EMA approaches by  
252 developing a novel and replicable methodology of standardized time-triggered smartphone-based  
253 EMAs to capture prospective PA within free-living social and physical contexts by leveraging  
254 citizen-owned smartphones running on both Android and iOS systems.

255 We were able to not only develop a novel EMA that can be time-triggered by both iOS and  
256 Android devices to capture prospective PA across physical and social contexts to address current  
257 gaps in EMA methodologies<sup>22,23</sup>, but also compared this EMA measure with IPAQ to highlight  
258 potential discrepancies between prospective and retrospective measures in capturing active living  
259 in free-living conditions.

260 Although not statistically significant, irrespective of the intensity of PA (overall PA, light,  
261 and moderate-to-vigorous PA), citizen scientists consistently over-reported activity using IPAQ.  
262 However, when PA intensities were compared across the 4 physical domains (workplace; active  
263 transportation; household; and recreation, sport, and leisure), PA reported via IPAQ in active  
264 transportation; and recreation, sport, and leisure domains was significantly greater than PA  
265 reported via EMAs.

1  
2  
3 266 These findings corroborated a longitudinal validation study by Swendeman et al., (2018),  
4  
5 267 who concluded that the inter-method reliability between smartphone-based EMAs and their  
6  
7  
8 268 corresponding recall reports was low and no significant associations were observed.<sup>24</sup> Another  
9  
10 269 validation study that compared PA EMAs with IPAQ and accelerometer measures concluded that  
11  
12 270 EMA measures correlated better with accelerometers.<sup>4</sup> Several studies have been conducted to  
13  
14 271 compare self-report estimates of PA with objective measures (an accelerometer),<sup>4,25,26</sup> with  
15  
16 272 evidence suggesting that an ideal approach potentially lies between self-reports and  
17  
18  
19 273 accelerometry,<sup>24</sup> especially because accelerometry is unable to capture context.

22 274 This is indicative of EMAs being the potential solution for comprehensively capturing PA by  
23  
24 275 minimizing recall bias. However, a key gap in current methodologies is that EMAs are used in  
25  
26 276 more controlled experiments, where identical mobile devices running on same operating systems  
27  
28  
29 277 are being provided to participants (Refs). Moreover, EMA methodologies lack standardization  
30  
31 278 and sufficient rigour. Another important gap is the inability of existing EMAs to capture critical  
32  
33 279 physical and social contexts within which PA is accumulated.

36 280 In our study we addressed these gaps by adopting a citizen science approach<sup>27</sup>, where  
37  
38 281 participants used their own smartphones, which operated on either iOS and Android systems,  
39  
40 282 thus expanding the scope of leveraging these ubiquitous tools<sup>28</sup> to conduct ethical surveillance<sup>8,29</sup>  
41  
42  
43 283 of PA among large populations. Citizen science approaches are increasingly being considered in  
44  
45 284 active living research<sup>30</sup>, and it is important that methodological advancements are in step with  
46  
47 285 conceptual and technological innovations. Another key advancement of this study is including  
48  
49 286 only those participants who completed EMAs on at least 3 days, an inclusion criterion which  
50  
51 287 provides the necessary rigour to arrive at valid data.



1  
2  
3 288 However, perhaps the most important addition to the methodology was introducing a looped  
4  
5 289 linking EMA that not only captured the intensity and volume of PA, but also the physical and  
6  
7  
8 290 social contexts of PA (**Figures 3 and 4**). The findings showed that citizen scientists reported  
9  
10 291 accumulating most PA while at home, through active transportation, at the gym, at their work  
11  
12 292 places, and in parks.<sup>15,31</sup> The distribution of accumulation of overall PA across these physical  
13  
14 293 contexts provides important evidence to develop interventions modifying physical spaces to  
15  
16 294 address physical inactivity.<sup>32-35</sup> Perhaps even more interesting were the results of social context,  
17  
18 295 where the findings showed that most citizen scientists accumulated PA by themselves<sup>31</sup>, which  
19  
20 296 points towards informing individual-level interventions that facilitate intrinsic motivation.<sup>36-38</sup>  
21  
22  
23

24 297 Current evidence clearly indicates that there is no gold standard in assessing prospective PA  
25  
26 298 using mobile EMAs, and this study advances a methodology that introduces conceptual and  
27  
28 299 technological advancement (citizen science approach utilizing citizen-owned devices functioning  
29  
30 300 on both iOS and Android systems), scientific rigour (stringent inclusion criteria for valid data),  
31  
32 301 and comprehensiveness of data collection (volume, intensity, and context). In working towards a  
33  
34 302 standardized EMA methodology future studies need to address the balance between burden and  
35  
36 303 compliance. Moreover, future studies could combine EMAs with objective measurement to  
37  
38 304 measure PA,<sup>39,40</sup> to concretely capture PA prospectively. Nevertheless, EMAs have the potential  
39  
40 305 to reliably record active living and could substitute accelerometers when needed.<sup>2</sup>  
41  
42  
43  
44  
45  
46  
47  
48

### 49 **Strengths and Limitations**

50 308 The primary strength of the study is the development of novel and replicable methodology to  
51  
52 309 capture prospective PA comprehensively from large populations using citizen-owned devices.  
53  
54

1  
2  
3 310 The main limitation is the small sample size after applying the inclusion criteria, however,  
4  
5 311 smaller sample sizes are not uncommon in smartphone-based EMA studies.<sup>41</sup> Nevertheless, to  
6  
7 312 capitalize on the citizen science approach, it is important consider innovative solutions such as  
8  
9 313 crowdsourcing<sup>42,43</sup> to engage large populations for the ethical active living surveillance.<sup>8</sup>  
10  
11  
12

## 13 314 **CONCLUSION**

16 315 With growth of smartphones projected to only magnify in the future<sup>16</sup>, these ubiquitous tools  
17  
18 316 can be leveraged via citizen science to capture accurate active living patterns of large populations  
19  
20 317 in free-living conditions through innovative EMAs. This citizen science methodology adapted  
21  
22 318 mobile EMAs to minimize recall bias and capture not only prospective PA, but also important  
23  
24 319 physical and social contexts within which individuals accumulate PA.  
25  
26  
27

## 28 320 **LIST OF ABBREVIATIONS**

31 321 EMAs: Ecological momentary assessments  
32  
33

34 322 PA: Physical activity  
35  
36

37 323 IPAQ: International physical activity questionnaire  
38  
39

40 324  
41  
42

43 325  
44  
45

## 46 326 **DECLARATIONS**

49 327 **Ethics approval and consent to participate:** All citizen scientists provided informed consent  
50  
51 328 through the app and confirmed their age ( $\geq 18$  years) before being recruited. Ethics approval was  
52  
53

329 obtained from the universities of Regina and Saskatchewan through a synchronized review  
330 protocol (REB # 2017-29).

331 **Consent for publication:** Not applicable

332 **Availability of data and materials:** The corresponding author will make the data available  
333 upon reasonable request.

334 **Competing interests:** Authors declare no conflict of interest.

335 **Funding information:** This study was funded by the Saskatchewan Health Research  
336 Foundation's Establishment Grant. This funding body had no role in study design; data  
337 collection, analysis and interpretation; or in writing the manuscript.

338 **Authors' contributions:** TRK contributed substantially to the study design, acquisition and  
339 interpretation of data, and writing the manuscript. LMC contributed substantially to the  
340 acquisition and interpretation of data, and writing the manuscript. All authors read and approved  
341 the final manuscript.

342 **Acknowledgments:** The authors acknowledge the entire DEPTH Lab team, and especially Ms  
343 Kayla Brodersen for her role in the successful implementation of the SMART Platform.

## 344 REFERENCES

- 345 1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin*  
346 *Psychol.* 2008;4:1-32.
- 347 2. Burke LE, Shiffman S, Music E, *et al.* Ecological momentary assessment in behavioral  
348 research: addressing technological and human participant challenges. *J Med Internet Res.*  
349 2017;19(3):e77.

- 1  
2  
3 350 3. Thomas JG, Bond DS, Ryder BA, *et al.* Ecological momentary assessment of  
4  
5 351 recommended postoperative eating and activity behaviors. *Surg Obes Relat Dis.* Mar-Apr  
6  
7 352 2011;7(2):206-212.  
8  
9  
10 353 4. Knell G, Gabriel KP. Ecological Momentary Assessment of Physical Activity: Validation  
11  
12 354 Study. Jul 18 2017;19(7):e253.  
13  
14  
15 355 5. Dunton GF. Ecological Momentary Assessment in Physical Activity Research. *Exerc*  
16  
17 356 *Sport Sci Rev.* Jan 2017;45(1):48-54.  
18  
19 357 6. Prince SA, Adamo KB, Hamel ME, *et al.* A comparison of direct versus self-report  
20  
21 358 measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr*  
22  
23 359 *Phys Act.* Nov 6 2008;5:56.  
24  
25  
26 360 7. Marszalek J, Morgulec-Adamowicz N, Rutkowska I, *et al.* Using Ecological Momentary  
27  
28 361 Assessment to Evaluate Current Physical Activity. *BioMed Res.* 2014;2014:9.  
29  
30  
31 362 8. Katapally TR, Bhawra J, Leatherdale ST, *et al.* The SMART study, a mobile health and  
32  
33 363 citizen science methodological platform for active living surveillance, integrated  
34  
35 364 knowledge translation, and policy interventions: Longitudinal Study. *JMIR Public Health*  
36  
37 365 *& Surveill.* 2018;4(1):e31.  
38  
39  
40 366 9. Kanning M, Schlicht W. Be active and become happy: an ecological momentary  
41  
42 367 assessment of physical activity and mood. *J Sport Exerc Psychol.* Apr 2010;32(2):253-  
43  
44 368 261.  
45  
46  
47 369 10. Conroy DE, Elavsky S, Hyde AL, *et al.* The dynamic nature of physical activity  
48  
49 370 intentions: a within-person perspective on intention-behavior coupling. *J Sport Exerc*  
50  
51 371 *Psychol.* Dec 2011;33(6):807-827.  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 372 11. Matthews CE, Moore SC, George SM, *et al.* Improving self-reports of active and  
4  
5 373 sedentary behaviors in large epidemiologic studies. *Exerc Sport Sci Rev.* Jul  
6  
7 374 2012;40(3):118-126.  
8  
9  
10 375 12. Adams SA, Matthews CE, Ebbeling CB, *et al.* The effect of social desirability and social  
11  
12 376 approval on self-reports of physical activity. *Am J Epidemiol.* Feb 15 2005;161(4):389-  
13  
14 377 398.  
15  
16  
17 378 13. Hekler EB, Buman MP. Validation of Physical Activity Tracking via Android  
18  
19 379 Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living  
20  
21 380 Validation Studies. Apr 15 2015;3(2):e36.  
22  
23  
24 381 14. Rodriguez VH, Medrano C, Plaza I, *et al.* Estimating the Physical Activity with  
25  
26 382 Smartphones: Analysis of the Device Position and Comparison with GT3X+ Actigraph.  
27  
28 383 2016; Cham.  
29  
30  
31 384 15. Bejarano CM, Cushing CC, Crick CJ. Does context predict psychological states and  
32  
33 385 activity? An ecological momentary assessment pilot study of adolescents. *Psychol Sport*  
34  
35 386 *Exerc.* 2019/03/01/ 2019;41:146-152.  
36  
37  
38 387 16. Katapally TR. The SMART Framework: Integration of Citizen Science, Community-  
39  
40 388 Based Participatory Research, and Systems Science for Population Health Science in the  
41  
42 389 Digital Age. *JMIR Mhealth Uhealth.* Aug 30 2019;7(8):e14056.  
43  
44  
45 390 17. Katapally TR. SMART: A Mobile Health and Citizen Science Platform. [Internet]. 2019;  
46  
47 391 <https://www.smartstudysask.com/copy-of-smart-youth>. Accessed November 22, 2019.  
48  
49 392 18. Silvertown J. A new dawn for citizen science. *Trends in ecology & evolution.*  
50  
51 393 2009;24(9):467-471.  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 394 19. Katapally TR, Bhawra J, Leatherdale ST, *et al.* The SMART Study, a Mobile Health and  
4  
5 395 Citizen Science Methodological Platform for Active Living Surveillance, Integrated  
6  
7 396 Knowledge Translation, and Policy Interventions: Longitudinal Study. *JMIR public*  
8  
9 397 *health and surveillance*. 2018;4(1):e31-e31.
- 10  
11  
12 398 20. Craig CL, Marshall AL, Sjostrom M, *et al.* International physical activity questionnaire:  
13  
14 399 12-country reliability and validity. *Med Sci Sports Exerc*. Aug 2003;35(8):1381-1395.
- 15  
16  
17 400 21. Sullivan R, Kinra S, Ekelund U, *et al.* Evaluation of the Indian Migration Study Physical  
18  
19 401 Activity Questionnaire (IMS-PAQ): a cross-sectional study. *Int J Behav Nutr Phys Act*.  
20  
21 402 Feb 9 2012;9:13.
- 22  
23  
24 403 22. Marszalek J, Morgulec-Adamowicz N, Rutkowska I, *et al.* Using ecological momentary  
25  
26 404 assessment to evaluate current physical activity. *Biomed Res Int*. 2014;2014:915172.
- 27  
28  
29 405 23. Bedard C, King-Dowling S. Understanding Environmental and Contextual Influences of  
30  
31 406 Physical Activity During First-Year University: The Feasibility of Using Ecological  
32  
33 407 Momentary Assessment in the MovingU Study. May 31 2017;3(2):e32.
- 34  
35  
36 408 24. Swendeman D, Comulada WS, Koussa M, *et al.* Longitudinal Validity and Reliability of  
37  
38 409 Brief Smartphone Self-Monitoring of Diet, Stress, and Physical Activity in a Diverse  
39  
40 410 Sample of Mothers. *JMIR mHealth and uHealth*. 2018;6(9):e176.
- 41  
42  
43 411 25. Troiano RP, Berrigan D, Dodd KW, *et al.* Physical activity in the United States measured  
44  
45 412 by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181.
- 46  
47  
48 413 26. Atienza AA, Moser RP, Perna F, *et al.* Self-reported and objectively measured activity  
49  
50 414 related to biomarkers using NHANES. *Med Sci Sports Exerc*. 2011;43(5):815-821.

- 1  
2  
3 415 27. Mittelstadt B, Benzler J, Engelmann L, *et al.* Is there a duty to participate in digital  
4  
5 416 epidemiology? *Life Sciences, Society and Policy*. 2018/05/09 2018;14(1):9.  
6  
7  
8 417 28. Hinckson E, Schneider M, Winter SJ, *et al.* Citizen science applied to building healthier  
9  
10 418 community environments: advancing the field through shared construct and measurement  
11  
12 419 development. *Int. J. Behav. Nutr. Phys. Act.* 2017;14(1):133.  
13  
14  
15 420 29. Liao Y, Intille SS, Dunton GF. Using ecological momentary assessment to understand  
16  
17 421 where and with whom adults' physical and sedentary activity occur. *Int J Behav Med*. Feb  
18  
19 422 2015;22(1):51-61.  
20  
21  
22 423 30. Dollman J. Social and Environmental Influences on Physical Activity Behaviours.  
23  
24 424 Multidisciplinary Digital Publishing Institute; 2018.  
25  
26 425 31. Benton JS, Anderson J, Hunter RF, *et al.* The effect of changing the built environment on  
27  
28 426 physical activity: a quantitative review of the risk of bias in natural experiments. *Int. J.*  
29  
30 427 *Behav. Nutr. Phys. Act.* 2016;13(1):107.  
31  
32  
33 428 32. Coghill C-L, Valaitis RK, Eyles JD. Built environment interventions aimed at improving  
34  
35 429 physical activity levels in rural Ontario health units: a descriptive qualitative study. *BMC*  
36  
37 430 *public health*. 2015;15(1):464.  
38  
39  
40 431 33. Wilkie S, Townshend T, Thompson E, *et al.* Restructuring the built environment to  
41  
42 432 change adult health behaviors: a scoping review integrated with behavior change  
43  
44 433 frameworks. *Cities & health*. 2018;2(2):198-211.  
45  
46  
47 434 34. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile  
48  
49 435 technology into psychosocial and health behaviour treatments. *British journal of health*  
50  
51 436 *psychology*. 2010;15(1):1-39.  
52  
53

- 1  
2  
3 437 35. Daugherty DA, Runyan JD, Steenbergh TA, *et al.* Smartphone delivery of a hope  
4  
5 438 intervention: Another way to flourish. *PloS one*. 2018;13(6):e0197930.  
6  
7 439 36. Nahum-Shani I, Smith SN, Spring BJ, *et al.* Just-in-time adaptive interventions (JITAs)  
8  
9 440 in mobile health: key components and design principles for ongoing health behavior  
10  
11 441 support. *Ann. Behav. Med.* 2017;52(6):446-462.  
12  
13 442 37. Zink J, Belcher BR, Dzibur E, *et al.* Association between self-reported and objective  
14  
15 443 activity levels by demographic factors: ecological momentary assessment study in  
16  
17 444 children. *JMIR mHealth and uHealth*. 2018;6(6):e150.  
18  
19 445 38. Dunton GF, Liao Y, Intille SS, *et al.* Investigating children's physical activity and  
20  
21 446 sedentary behavior using ecological momentary assessment with mobile phones. *Obesity*.  
22  
23 447 2011;19(6):1205-1212.  
24  
25 448 39. Prince SA, Adamo KB, Hamel ME, *et al.* A comparison of direct versus self-report  
26  
27 449 measures for assessing physical activity in adults: a systematic review. *Int. J. Behav.*  
28  
29 450 *Nutr. Phys. Act.* 2008;5(1):56.  
30  
31 451 40. Kraemer JD, Strasser AA, Lindblom EN, *et al.* Crowdsourced data collection for public  
32  
33 452 health: A comparison with nationally representative, population tobacco use data. *Prev*  
34  
35 453 *Med.* Sep 2017;102:93-99.  
36  
37  
38  
39  
40  
41  
42

43 **Table 1. Overall physical activity measurement: IPAQ vs. EMA**  
44  
45

	Mean (minutes/day)	Standard Deviation	Percentiles (minutes/day)			p- value*
			25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	
IPAQ	194.39	266.10	63.80	122.14	175.72	0.331



EMA	140.91	98.31	73.07	123.75	183.48	

455 Note: Based on Wilcoxon Signed Ranks Test; IPAQ: international physical activity  
456 questionnaire; EMA: ecological momentary assessment

457

458

459 **Table 2. Light and moderate-to-vigorous physical activity measurement: IPAQ vs. EMA**

Intensity		Mean	Standard	Percentiles			p-value*
		(minutes/day)	Deviation	(minutes/day)			
				25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	
Light	IPAQ	116.99	171.24	36.00	67.86	110.00	0.322
	EMA	87.16	64.44	41.25	68.33	103.67	
Moderate to vigorous	IPAQ	98.42	175.18	17.14	49.44	92.86	0.995
	EMA	70.38	63.48	40.00	52.50	87.50	

460 Note: Based on Wilcoxon Signed Ranks Test; IPAQ: international physical activity  
461 questionnaire; EMA: ecological momentary assessment

462

463

464

465 **Table 3. Overall physical activity measurement across physical domains: IPAQ vs. EMA**

Domain		Mean	SD	Percentiles (minutes/day)			p-value*
		(minutes/day)		25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	
Workplace	Survey	32.14	34.97	7.86	12.86	66.07	0.345
	EMA	20.50	17.87	5.63	15.00	38.13	
Active Transportation PA	Survey	43.97	24.32	25.36	40.00	66.43	0.002
	EMA	16.41	11.51	7.56	10.00	25.69	
Household	Survey	38.27	35.01	9.04	28.50	74.46	0.117
	EMA	25.33	46.29	5.16	10.31	19.84	
Recreation Sport and Leisure-time PA	Survey	145.90	306.95	15.00	34.29	72.86	0.003
	EMA	20.88	16.19	7.50	17.50	37.50	

1  
2  
3  
4  
5 466 Note: Based on related-samples Wilcoxon Signed Rank Tests; IPAQ: international physical  
6 467 activity questionnaire; EMA: ecological momentary assessment  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

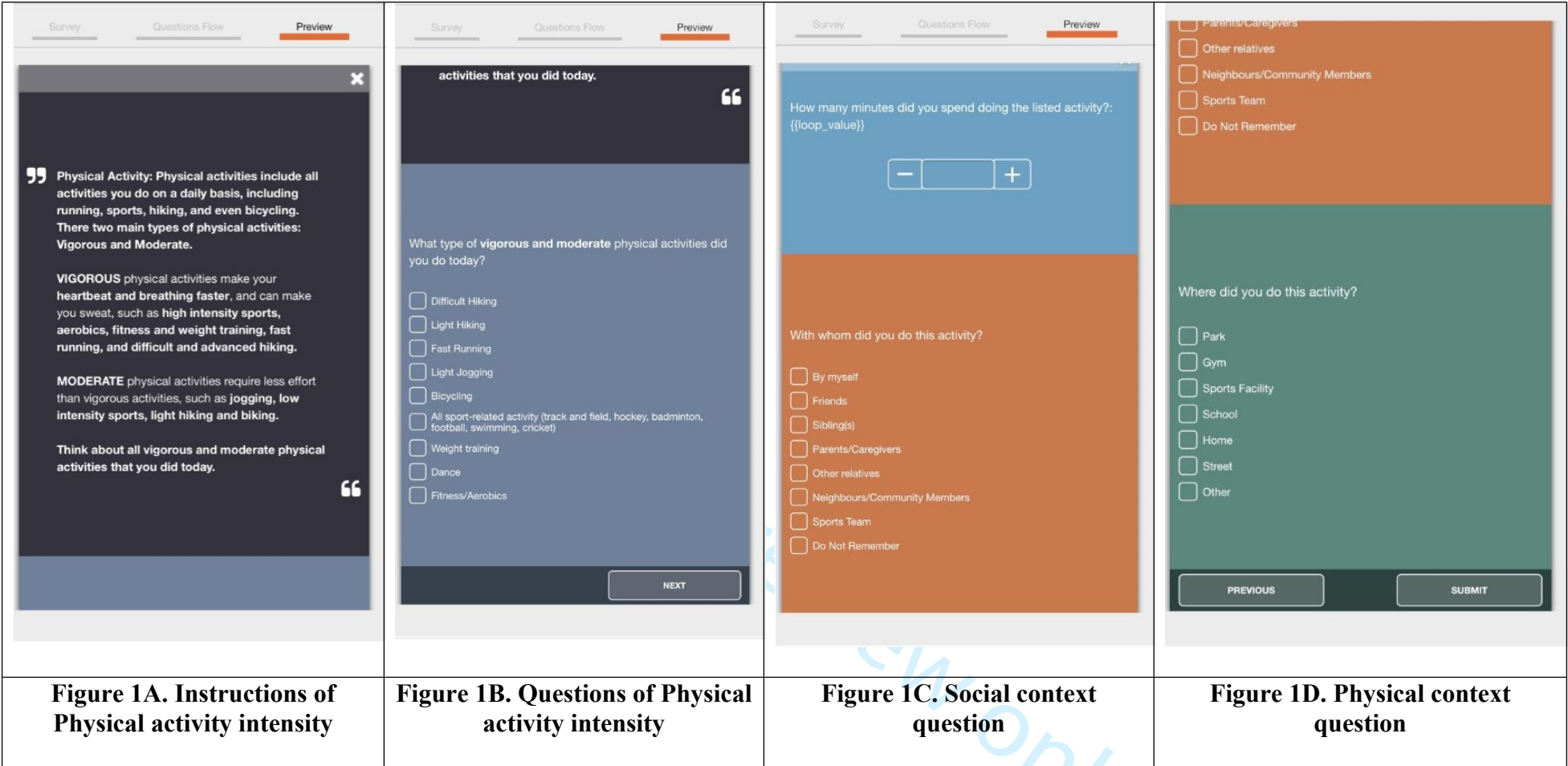


Figure 1. Time-triggered ecological momentary assessment capturing prospective physical activity

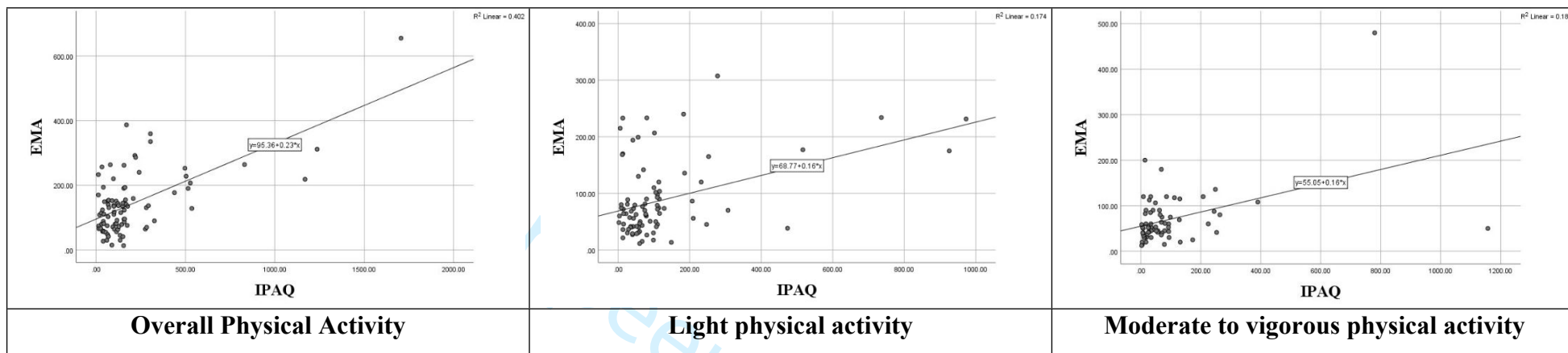


Figure 2. Correlation between IPAQ and EMA measurements of physical activity

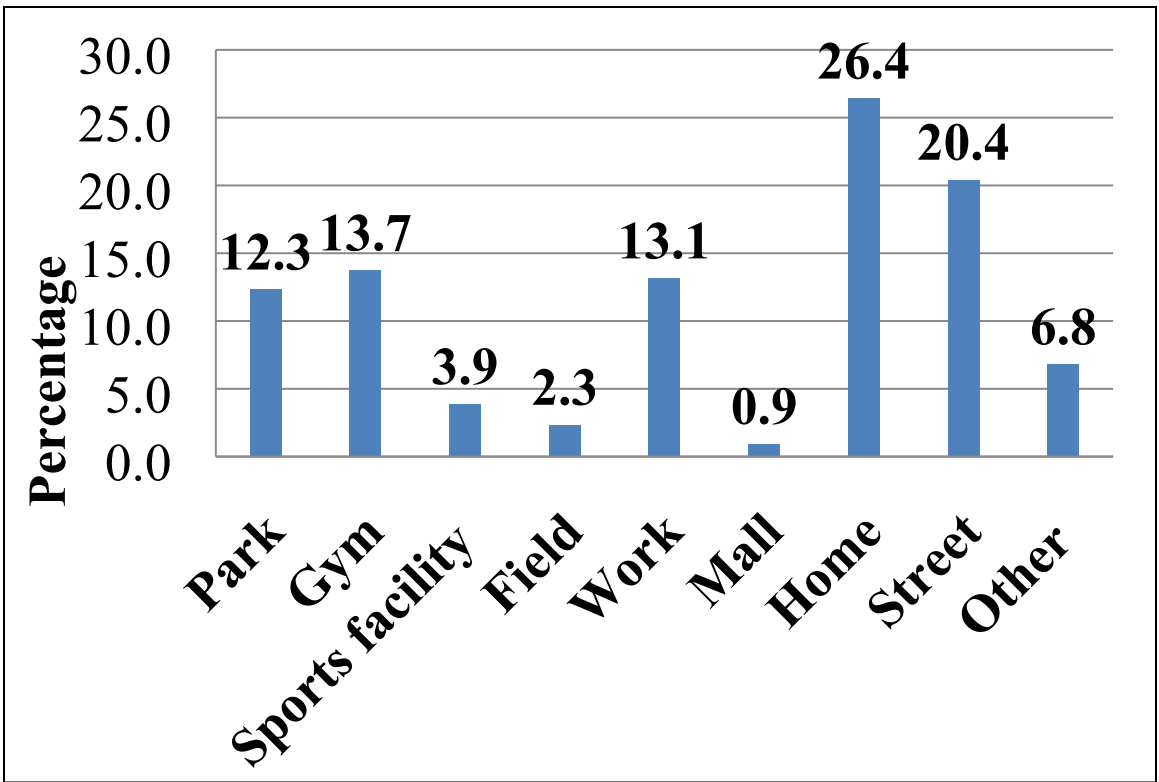


Figure 3. Distribution of Daily EMA Physical activity within physical contexts

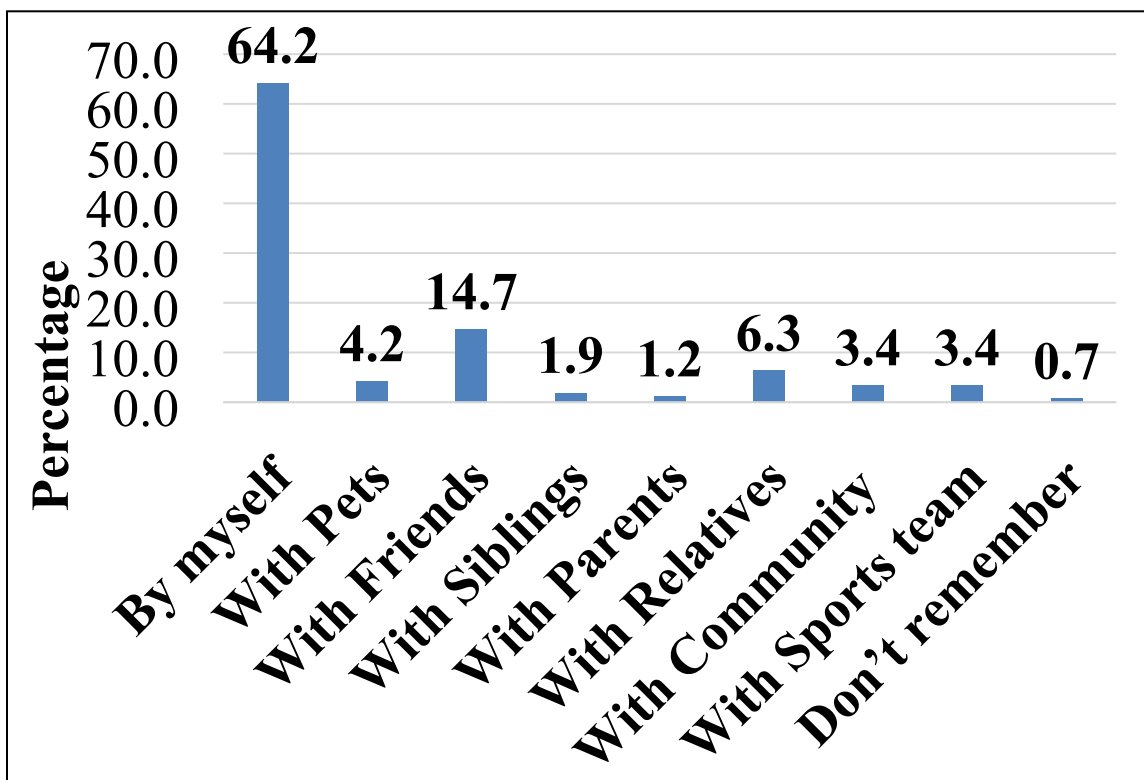


Figure 4. Distribution of Daily EMA Physical activity within social contexts

## STROBE Statement

Checklist of items that should be included in reports of observational studies

Section/Topic	Item No	Recommendation	Reported on Page No
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
<b>Introduction</b>			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
<b>Methods</b>			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6
		<i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	
		<i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
Variables	7	(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed	6,7,8,9
		<i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Data sources/measurement	8*	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6,7,8,9
Bias	9	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6,7
Study size	10	Describe any efforts to address potential sources of bias	9
Quantitative variables	11	Explain how the study size was arrived at	9
Statistical methods	12	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6,7,8,9
		(a) Describe all statistical methods, including those used to control for confounding	9
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	9
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed	
		<i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed	9
		<i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	

Section/Topic	Item No	Recommendation	Reported on Page No
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	10, 11
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10,11
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key results	18	Summarise key results with reference to study objectives	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12.13
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).



# BMJ Open

## A digital epidemiological and citizen science methodology to capture prospective physical activity in free-living conditions: A SMART platform study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-036787.R1
Article Type:	Original research
Date Submitted by the Author:	24-Apr-2020
Complete List of Authors:	Katapally , Tarun ; University of Regina, ; University of Saskatchewan, Chu, Luan; University of Saskatchewan,
<b>Primary Subject Heading</b>:	Epidemiology
Secondary Subject Heading:	Public health
Keywords:	EPIDEMIOLOGY, PUBLIC HEALTH, SPORTS MEDICINE

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3 **Title:** A digital epidemiological and citizen science methodology to capture prospective physical  
4 activity in free-living conditions: A SMART platform study  
5  
6  
7

8 Authors: Tarun Reddy Katapally, MS, PhD,<sup>1,2,3</sup>, Luan Manh Chu, MS<sup>3,4</sup>  
9  
10

11 <sup>1</sup>Johnson Shoyama Graduate School of Public Policy, University of Regina, 2155 College Ave,  
12 Regina, SK S4M0A1, Canada. Tarun.katapally@uregina.ca.  
13

14 <sup>2</sup>Johnson Shoyama Graduate School of Public Policy, University of Saskatchewan, 101  
15 Diefenbaker Pl, Saskatoon, SK S7N 5B8, Canada. Tarun.katapally@usask.ca.  
16  
17

18 <sup>3</sup>College of Medicine, Health Science Building, 107 Wiggins Road, University of Saskatchewan,  
19 Saskatoon, SK S7N 5E5, Canada. Tarun.katapally@usask.ca.  
20  
21

22 <sup>4</sup>Canadian Centre for Health and Safety in Agriculture, University of Saskatchewan, 104 Clinic  
23 Place, PO Box 23, Saskatoon, SK S7N 2Z4, Canada. Cml779@mail.usask.ca  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

42 Corresponding author: Tarun Reddy Katapally  
43 Johnson Shoyama Graduate School of Public Policy, University of Regina  
44 110-2 Research Drive, Innovation Place, Regina, SK, S4S7H9, Canada  
45 Phone: 1 3065854544  
46 Email: [Tarun.katapally@uregina.ca](mailto:Tarun.katapally@uregina.ca)  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 37 **ABSTRACT**  
4

5 38 **Objectives** The purpose of this study was to develop a replicable methodology of mobile  
6  
7 39 ecological momentary assessments (EMAs) to capture prospective physical activity (PA) within  
8  
9 40 free-living social and physical contexts by leveraging citizen-owned smartphones running on  
10  
11 41 both Android and iOS systems.  
12

13  
14 42 **Design** Data were obtained from the cross-sectional pilots of the SMART Platform, a citizen  
15  
16 43 science and mobile health initiative.  
17

18  
19 44 **Setting** The cities of Regina and Saskatoon, Canada.  
20

21 45 **Participants** 538 citizen scientists ( $\geq 18$  years) provided PA data during 8 consecutive days using  
22  
23 46 a custom-built smartphone app, and after applying a rigid inclusion criteria, 89 were included in  
24  
25 47 the final analysis.  
26

27  
28 48 **Outcome measures** EMAs enabled reporting of light, moderate, and vigorous PA, as well as  
29  
30 49 physical and social contexts of PA. Retrospective PA was reported using International Physical  
31  
32 50 Activity Questionnaire (IPAQ). For both measures, PA intensities were categorized into mean  
33  
34 51 minutes of light and moderate-to-vigorous PA per day. Wilcoxon signed ranks tests and  
35  
36 52 Spearman correlation procedures were conducted to compare PA intensities reported via EMAs  
37  
38 53 and IPAQ.  
39

40  
41 42 **Results** Using EMAs, citizen scientists reported 140.91, 87.16, and 70.38 mean minutes/day of  
43  
44 45 overall, light, and moderate-to-vigorous PA, respectively, whereas using IPAQ they reported  
46  
47 48 194.39, 116.99, and 98.42 mean minutes/day of overall, light, and moderate-to-vigorous PA,  
49  
50 49 respectively. Overall ( $\rho=0.414$ ,  $p<0.001$ ), light ( $\rho =0.261$ ,  $p=0.012$ ) and moderate-to-vigorous  
51  
52 58 PA ( $\rho =0.316$ ,  $p=0.009$ ) were fairly correlated between EMA and IPAQ. In comparison with  
53

EMAs, using IPAQ citizen scientists reported significantly greater overall PA in active transportation ( $p=0.002$ ) and recreation, sport and leisure-time domains ( $p=0.003$ ).

**Conclusions** This digital epidemiological and citizen science methodology adapted mobile EMAs to capture not only prospective PA, but also important physical and social contexts within which individuals accumulate PA. Ubiquitous tools can be leveraged via citizen science to capture accurate active living patterns of large populations in free-living conditions through innovative EMAs.

**Keywords:** Physical activity, mHealth, Ecological Momentary Assessments, Measurement, Citizen Science, Digital Epidemiology

### Strengths

- The methodology addresses current discrepancies in mobile ecological momentary assessments (EMAs) (e.g., triggering processes, time to follow-up).
- The methodology of time-triggered mobile EMAs is effective in recording comprehensive daily prospective physical activity.
- The methodology facilitates capture of both physical and social contexts of physical activity prospectively.

### Limitations

- The main limitation is the small sample size after applying the inclusion criteria
- All observations are self-reported by citizen scientists.

## 84 INTRODUCTION

85 Advances in mobile technology over the past decade have facilitated the innovation of  
86 ecological momentary assessments (EMAs), which are digital epidemiological tools that aid in  
87 understanding environmental, social, and behavioural processes.<sup>1,2</sup> EMAs can capture real-time  
88 data that reflect the dynamics of participants' experiences in their natural environment and thus  
89 they are increasingly being used to monitor health behaviors among populations across the life  
90 course.<sup>3-5</sup> In active living research, evidence indicates that EMAs are a valid, reliable, and  
91 feasible method of data collection.<sup>6,7</sup>

92 EMAs are an advancement over traditional self-report methods as they enable data collection  
93 more proximal to the time and place that a behavior has occurred.<sup>2,8</sup> Moreover, EMAs overcome  
94 many of the limitations of traditional self-report surveys to provide information regarding  
95 specific activity types (e.g. watching TV vs video gaming) and capture important factors that  
96 influence health behaviors such as mood and environmental perceptions.<sup>5,9,10</sup> In measuring  
97 physical activity (PA) intensities, EMAs have been shown to minimize recall<sup>6,11</sup> and social  
98 desirability bias<sup>12</sup> of traditional self-report measures.

99 Several studies have examined the validity of smartphone-based EMAs compared to other  
100 objective devices (accelerometers, pedometers) and self-report measures of PA.<sup>2,10-16</sup> Overall,  
101 estimates from EMAs were found to be highly correlated with accelerometer estimates.<sup>13,14</sup>  
102 However, this evidence also indicates that PA was over-reported when International Physical  
103 Activity Questionnaire (IPAQ) was used<sup>2</sup> and that daily PA EMA reports were not significantly  
104 associated with their traditional recall measures.<sup>12</sup>

1  
2  
3 105 Currently there is little evidence of existing EMA methods that capture PA intensities across  
4  
5 106 various physical (leisure-time PA, transit-related PA, occupation-related PA, and  
6  
7 107 household/domestic-related PA)<sup>15</sup> and social contexts (with family, friends etc.).<sup>16</sup> Moreover,  
8  
9 108 there are discrepancies in smartphone-based EMA methodologies, which range from inconsistent  
10  
11 109 EMA triggering processes and varying times of prospective follow-up, to limitations of using  
12  
13 110 identical mobile devices and operating systems.<sup>8</sup>  
14  
15  
16  
17

18 111 The objective of this study is to address current deficiencies in active living EMA approaches  
19  
20 112 by developing a replicable digital epidemiological and citizen science methodology to capture  
21  
22 113 prospective PA within free-living social and physical contexts. This objective will be achieved  
23  
24 114 by leveraging citizen-owned smartphones running on both Android and iOS systems, and by  
25  
26 115 comparing EMA measures with traditional self-report measures of PA within the same cohort.  
27  
28  
29

## 30 116 **METHODS**

### 31 32 33 117 **Design**

34  
35  
36 118 This study is part of the SMART Platform, which is a mobile health (mHealth) and citizen  
37  
38 119 science initiative for active living surveillance, integrated knowledge translation, and policy and  
39  
40 120 real-time interventions.<sup>8,17,18</sup> Citizen science is a participatory approach where participants,  
41  
42 121 termed citizen scientists, actively engage in the research process from data collection to  
43  
44 122 knowledge translation, thus improving the probability of longitudinal participant compliance.<sup>15</sup> A  
45  
46 123 detailed description of SMART Platform's methods, including recruitment and data collection  
47  
48 124 strategies, are described in the Platform's methodology publication.<sup>8</sup>  
49  
50  
51  
52  
53

1  
2  
3 125 The data for this study have been obtained from the 2017 (April 1 to May 31) and 2018  
4  
5 126 (January 4 to March 31) cohorts of the SMART Platform,<sup>8</sup> which is a prospective investigation  
6  
7 127 designed to capture active living data from adults residing in the two largest urban centers in  
8  
9 128 Saskatchewan, Canada (Regina and Saskatoon). All subjective (via traditional validated surveys  
10  
11 129 and EMAs) and objective data (via smartphones sensors) related to PA, sedentary behaviour, and  
12  
13 130 perception of environment, individual motivation, health outcomes, and eudaimonic well-being  
14  
15 131 were obtained through citizen-owned smartphones on 8 consecutive days (**Figure 1**).  
16  
17  
18  
19

### 20 132 **Patient and public involvement**

21  
22 133 Participants in the SMART Platform are “citizen scientists” as they can engage with the  
23  
24 134 researchers at all stages of the research process. Thus, citizen scientists informed the design,  
25  
26 135 research questions and outcome measures. As part of the social media campaign for recruitment,  
27  
28 136 citizen scientists were encouraged to inform their friends about the study. Finally, as integrated  
29  
30 137 knowledge translation is part of the SMART Platform, results are disseminated throughout the  
31  
32 138 study period using the community voices webpage of the Platform’s website:  
33  
34 139 <https://www.smartstudysask.com/community-voices>  
35  
36  
37  
38

### 39 140 **Recruitment and participants**

40  
41 141 Citizen scientists for SMART Adult cohorts were recruited online through social media, and  
42  
43 142 in-person from the universities of Regina and Saskatchewan and community centres located in  
44  
45 143 different neighbourhoods in each city to capture a socioeconomically representative sample.  
46  
47 144 Citizen scientists were guided to download Ethica (Ethica Data Services Inc.), an  
48  
49 145 epidemiological smartphone application (app), specifically adapted for the SMART Platform,  
50  
51 146 which captures data through both Android and iOS platforms. All citizen scientists provided  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 147 informed consent through the app and confirmed their age ( $\geq 18$  years) before joining the study.  
4  
5 148 Ethics approval was obtained from the universities of Regina and Saskatchewan through a  
6  
7  
8 149 synchronized review protocol (REB # 2017-29).  
9

## 10 150 **Measures**

11  
12  
13 151 The two primary measures used in this study are the International Physical Activity  
14  
15 152 Questionnaire (IPAQ),<sup>19</sup> which collects retrospective PA in 4 physical domains (recreation,  
16  
17 153 active transportation, work, and home), and the SMART Platform's modified EMA, which  
18  
19  
20 154 captures prospective daily PA in both social and physical contexts.  
21  
22

### 23 155 *IPAQ*

24  
25  
26 156 IPAQ was deployed at baseline as soon as citizen scientists downloaded the app to self-report  
27  
28 157 physical activities over the past 7 days that were of at least 10 consecutive minutes in duration.  
29  
30 158 These activities were categorized by 4 domains: 1) Recreation (e.g., weight training, sports  
31  
32 159 (soccer, hockey, etc.), aerobics, running, jogging, swimming, cycling, etc.); 2) Household (e.g.,  
33  
34 160 carrying light loads, sweeping, washing windows, and raking, etc.); 3) Transportation (e.g.,  
35  
36 161 travelling in a train, bus, car, or other kind of motor vehicle, etc.); and 4) Work (e.g. heavy  
37  
38 162 lifting, digging, heavy construction or climbing upstairs, etc.). The records included the number  
39  
40  
41  
42 163 of times per week (within the last 7 days) and average minutes per day for each activity.  
43  
44

### 45 164 *Adapted Daily EMAs*

46  
47  
48 165 Using the SMART platform, time-triggered modified EMAs (**Figure 1A-D**) were developed,  
49  
50 166 tested, and piloted, before being pushed to citizen scientists' smartphones between 8pm and  
51  
52 167 8:30pm on each day for 8 consecutive days. These EMAs were designed to expire at 3am the  
53  
54

1  
2  
3 168 next day. Citizen scientists were asked to report only those physical activities that were of at  
4  
5 169 least 10 minutes in duration at a time. More importantly, each EMA was designed to not only  
6  
7 170 measure intensity and volume (in minutes) of PA, but also to capture social (i.e., with whom they  
8  
9 171 accumulated PA [**Figure 1C**]) and physical contexts (i.e., where they accumulated PA [**Figure**  
10  
11 172 **1D**]). This design was achieved by creating a looped linkage, where upon entering the type and  
12  
13 173 volume of each activity, the EMA triggered the social and physical context questions.  
14  
15  
16  
17

## 18 174 **Derived variables – Intensities and Volume of PA**

### 19 20 21 175 *IPAQ*

22  
23  
24 176 Thirty-seven questions related to PA were asked and 3 different categories of intensities were  
25  
26 177 created (light, moderate, and vigorous PA) by combining PA across 4 domains: recreation,  
27  
28 178 household, workplace, and active transportation. Moderate and vigorous PA intensities are  
29  
30 179 combined to derive “moderate-to-vigorous PA.” After conducting several aggregation  
31  
32 180 techniques, 2 final intensity variables were derived for IPAQ retrospective PA: mean minutes per  
33  
34 181 day of light and moderate-to-vigorous PA.  
35  
36  
37

### 38 182 *Adapted Daily EMAs*

39  
40  
41 183 A similar approach was employed to derive two final intensity variables for EMA prospective  
42  
43 184 PA: mean minutes per day of light and moderate-to-vigorous PA. For example, the light PA  
44  
45 185 included walking, light hiking, any light physical activity/sport (e.g. golf bowling etc.), yoga, and  
46  
47 186 light intensity household chores (e.g. washing dishes sweeping laundry gardening). Moderate-to-  
48  
49 187 vigorous PA included moderate to vigorous hiking, running, biking, any team sport (football  
50  
51 188 hockey soccer etc.), any other sport or activity (swimming canoeing skiing etc.), weight training,  
52  
53  
54

1  
2  
3 189 dance/aerobic/cardio exercise, and moderate-to-vigorous intensity household chores (e.g.  
4  
5 190 shovelling driveways, washing a car etc.).  
6  
7

### 8 191 *Physical Context*

10  
11 192 PA information from the IPAQ and EMAs (based on the question "Where did you do this  
12  
13 activity?") were grouped into domains. Domain 1: PA at workplace (IPAQ) and from work  
14 193 (EMAs). Domain 2: Transportation PA (IPAQ) and from street (EMAs). Domain 3: Housework,  
15  
16 194 house maintenance, and caring from family (IPAQ) and from home (EMAs). Domain 4:  
17  
18 195 Recreation, sport, and leisure-time PA (IPAQ) and from park, gym, and sport facility (EMAs).  
19  
20  
21 196  
22

### 23 197 *Social Context*

24  
25  
26 198 Social context information was collected via EMA question, "With whom did you do this  
27  
28 activity?" for each physical activity that the participants reported. Categories for social context  
29 199 included "by myself, with my dog, with my friend(s), with my parent(s)," among others.  
30  
31 200  
32  
33

### 34 201 *Statistical analyses*

35  
36  
37 202 The inclusion criterion to determine the final sample was dependent on citizen scientists  
38  
39 203 completing the IPAQ, and answering the daily EMA on at least 3 days. Continuous estimates  
40  
41 204 were reported as means with standard deviations (SD) and medians with 25<sup>th</sup> and 75<sup>th</sup> percentiles,  
42  
43 205 depending on normality. Where estimates were non-normal and positively skewed, median and  
44  
45 206 interquartile ranges were used. Wilcoxon signed ranks tests and Spearman correlation procedures  
46  
47 207 were conducted to compare PA intensities and domain-based PA reported via IPAQ and EMAs.  
48  
49  
50 208 Correlation coefficient values of <0.20, 0.21-0.40, 0.41-0.60, 0.61-0.80 and 0.81-1.0 were  
51  
52  
53

1  
2  
3 209 considered as weak, fair, moderate, strong and very strong correlation, respectively.<sup>20</sup>. Analyses  
4  
5 210 were conducted in SPSS version 24.0 (SPSS Inc., Chicago IL, USA) with significance set at  
6  
7  
8 211  $\alpha < 0.05$ .

## 11 212 **RESULTS**

13  
14 213 After applying the decision rule of including only those citizen scientists who completed  
15  
16 214 IPAQ, and answered the daily EMA on at least 3 days, out of 538 participants, only 89 were  
17  
18 215 included in this study (Table 1), among whom 47 identified as female (51.68%), and 26  
19  
20 216 identified as male (29.21%), and 19.11% (n=17) did not reveal their identity. The final sample  
21  
22 217 had the mean age of 37.15 years (SD=15.92), and a mean body mass index of 28.46 (SD=7.78).  
23  
24 218 The median (25<sup>th</sup>, 75<sup>th</sup> percentiles) and the mean (SD) duration of time (minutes per day) spent in  
25  
26 219 each of the activity intensities (light, moderate and vigorous), as well as overall PA were derived  
27  
28 220 from both IPAQ and EMA measures.

29  
30  
31 221 Using EMAs, citizen scientists reported 140.91, 87.16, and 70.38 mean minutes/day of  
32  
33 222 overall PA, light PA, and moderate-to-vigorous PA. The same citizen scientists reported 194.39,  
34  
35 223 116.99, and 98.42 mean minutes/day of overall PA, light PA, and moderate-to-vigorous PA  
36  
37 224 using the IPAQ (Table 2). These findings show that although there are no significant differences  
38  
39 225 between activity intensities reported via EMAs and IPAQ, citizen scientists consistently  
40  
41 226 overestimated their PA using IPAQ in comparison with EMAs. Table 3 demonstrates the  
42  
43 227 correlation between EMA and IPAQ measures to show that overall ( $\rho=0.414$ ,  $p<0.001$ ), light ( $\rho$   
44  
45 228  $=0.261$ ,  $p=0.012$ ) and moderate-to-vigorous PA ( $\rho =0.316$ ,  $p=0.009$ ) were fairly correlated across  
46  
47  
48  
49  
50 229 both measures. Figure 2 shows the visual representation of these correlations.

1  
2  
3 230 Figures 3 and 4 demonstrate the distribution of overall PA accumulated across different  
4  
5 231 physical and social contexts, as reported by citizen scientists using EMAs. Among physical  
6  
7 232 contexts, citizen scientists reported accumulating overall PA predominantly at home (26.4%), on  
8  
9 233 the streets ([20.4%] i.e., active transportation), at the gym (13.7%), at work (13.1%), and in parks  
10  
11 234 (12.3%). When it comes to social context, citizen scientists overwhelmingly reported  
12  
13 235 accumulating overall PA by themselves (64.2%), with some reporting being active with friends  
14  
15 236 (14.7%) and relatives (6.3%).

16  
17  
18  
19 237 As IPAQ captures PA in 4 physical domains (workplace, active transportation, household,  
20  
21 238 and recreation, sport and leisure-time) to compare estimates between EMA and IPAQ, EMA  
22  
23 239 estimates of overall PA accumulated across various physical contexts were categorized to match  
24  
25 240 the physical domains of IPAQ. Using EMAs, citizen scientists reported 20.50, 16.41, 25.33, and  
26  
27 241 20.88 mean minutes/day of overall PA across workplace, active transportation, household, and  
28  
29 242 recreation, sport and leisure-time domains, respectively. Using IPAQ, the same citizen scientists  
30  
31 243 reported 32.14, 43.97, 38.27, and 145.90 mean minutes/day of overall PA across workplace,  
32  
33 244 active transportation, household, and recreation, sport and leisure-time domains, respectively.  
34  
35  
36  
37

38 245 These findings show that in comparison with EMAs, there is a consistent pattern of over-  
39  
40 246 reporting of overall PA across all physical domains when citizen scientists used IPAQ, with  
41  
42 247 statistically significant differences observed in active transportation ( $p=0.002$ ) and recreation,  
43  
44 248 sport and leisure-time domains ( $p=0.003$ ) (Table 4). Table 5 demonstrates correlation between  
45  
46 249 EMA and IPAQ measures for overall physical activity across four physical domains, with  
47  
48 250 moderate correlation being depicted in household ( $\rho =0.607$ ,  $p=0.036$ ), and recreation, sport, and  
49  
50 251 leisure-time domains ( $\rho =0.587$ ,  $p=0.021$ ) domains.  
51  
52  
53

## 252 **DISCUSSION**

253 The objective of this study was to address current deficiencies in PA EMA approaches by  
254 developing a novel and replicable methodology of standardized time-triggered smartphone-based  
255 EMAs to capture prospective PA within free-living social and physical contexts by leveraging  
256 citizen-owned smartphones running on both Android and iOS systems.

257 We were able to not only develop a novel EMA that can be time-triggered by both iOS and  
258 Android devices to capture prospective PA across physical and social contexts to address current  
259 gaps in EMA methodologies<sup>21,22</sup>, but also compared this EMA measure with IPAQ to highlight  
260 potential discrepancies between prospective and retrospective measures in capturing active living  
261 in free-living conditions.

262 Although not statistically significant, irrespective of the intensity of PA (overall PA, light,  
263 and moderate-to-vigorous PA), citizen scientists consistently over-reported activity with IPAQ in  
264 comparison with EMA. However, when PA intensities were compared across the 4 physical  
265 domains (workplace; active transportation; household; and recreation, sport, and leisure), PA  
266 reported via IPAQ in active transportation; and recreation, sport, and leisure domains was  
267 significantly greater than PA reported via EMAs.

268 These findings corroborated a longitudinal validation study by Swendeman et al., (2018),  
269 who concluded that the inter-method reliability between smartphone-based EMAs and their  
270 corresponding recall reports was low and no significant associations were observed.<sup>23</sup> Another  
271 validation study that compared PA EMAs with IPAQ and accelerometer measures concluded that  
272 EMA measures correlated better with accelerometers.<sup>4</sup> Several studies have been conducted to

1  
2  
3 273 compare self-report estimates of PA with objective measures (an accelerometer),<sup>4,24,25</sup> with  
4  
5 274 evidence suggesting that an ideal approach potentially lies between traditional validated self-  
6  
7  
8 275 report measures and accelerometry,<sup>24</sup> especially because accelerometry is unable to capture  
9  
10 276 context.

11  
12  
13 277 This is indicative of EMAs being the potential solution for comprehensively capturing PA by  
14  
15 278 minimizing recall bias. However, a key gap in current methodologies is that EMAs are used in  
16  
17 279 more controlled experiments, where identical mobile devices running on same operating systems  
18  
19  
20 280 are to participants.<sup>26</sup> Moreover, EMA methodologies lack standardization and sufficient rigour  
21  
22 281 such as inclusion criteria for valid data. A key advancement of our study is including only those  
23  
24 282 participants who completed EMAs on at least 3 days, an inclusion criterion which provides the  
25  
26  
27 283 necessary rigour to arrive at valid data.

28  
29 284 EMAs are currently novel methods that are in need of standardization. We applied a strict  
30  
31 285 inclusion criterion, where we included only participants with PA data on at least 3 out of 8 days  
32  
33 286 in the final analysis, which resulted in exclusion of most participants. We did this even at the risk  
34  
35  
36 287 of reducing our sample size because this rigorous inclusion criterion is an essential step in  
37  
38 288 standardizing EMA measures, and obtaining valid and reliable data. This is not very different  
39  
40 289 from accelerometry standardization methods, where data are considered valid if participants wear  
41  
42  
43 290 accelerometers for at least several hours (e.g. 10 hours) on at least 2-3 days in a one-week study  
44  
45 291 period.<sup>27,28</sup>

46  
47 292 Another gap in current methodologies is the inability of existing EMAs to capture important  
48  
49 293 physical and social contexts within which PA is accumulated. We developed an innovative  
50  
51  
52 294 looped linking mechanism that sequentially triggers questions about type, volume, and context of

1  
2  
3 295 PA. The findings showed that citizen scientists reported accumulating most PA while at home,  
4  
5 296 through active transportation, at the gym, at their work places, and in parks.<sup>16,29</sup> The distribution  
6  
7  
8 297 of accumulation of overall PA across these physical contexts provides important evidence to  
9  
10 298 develop interventions modifying physical spaces to address physical inactivity.<sup>30-33</sup> The findings  
11  
12 299 also showed that most citizen scientists accumulated PA by themselves<sup>29</sup>, which points towards  
13  
14 300 informing individual-level interventions that facilitate intrinsic motivation.<sup>34-36</sup> Although these  
15  
16 301 findings are not novel by themselves, the methodology of using a single time-triggered EMA per  
17  
18 302 day to capture volume, intensity, and physical and social contexts of PA is innovative.

19  
20  
21 303 Although EMAs are valid and reliable measures to measure PA, current evidence indicates  
22  
23 304 that there is no gold standard in assessing prospective PA using mobile EMAs.<sup>3,5,21</sup> Our study  
24  
25 305 advances a methodology that introduces conceptual and technological advancement (citizen  
26  
27 306 science approach utilizing citizen-owned devices functioning on both iOS and Android systems),  
28  
29 307 scientific rigour (stringent inclusion criteria for valid data), and comprehensiveness of data  
30  
31 308 collection (volume, intensity, and contexts). In working towards standardized EMA  
32  
33 309 methodology, future studies need to address the balance between capture of prospective PA and  
34  
35 310 participant burden/compliance in repeatedly responding to EMAs. Future studies should could  
36  
37 311 combine EMAs with objective measurement to measure PA,<sup>37,38</sup> to concretely capture PA.

38  
39  
40  
41  
42 312 Nevertheless, EMAs for PA measurement have the potential to reliably record active living  
43  
44 313 and could substitute accelerometers when needed.<sup>2</sup> In our study we addressed existing gaps in  
45  
46 314 EMA methodology to measure PA by adopting a citizen science approach<sup>39</sup> in deploying a  
47  
48 315 comprehensive, yet generic EMA that captures type, volume, and context of PA. More  
49  
50 316 importantly, participants used their own smartphones, which operated on either iOS or Android



1  
2  
3 317 systems. Thus, this methodology is not only replicable, but also expands the scope of leveraging  
4  
5 318 ubiquitous tools such as smartphones<sup>40</sup> to conduct ethical surveillance<sup>8,41</sup> of PA among large  
6  
7 319 populations. Citizen science approaches are increasingly being considered in active living  
8  
9 320 research<sup>42</sup>, and it is important that methodological advancements are in step with conceptual and  
10  
11 321 technological innovations. With more than 3 billion smartphones currently in circulation  
12  
13 322 globally,<sup>40</sup> standardized and generic EMA methodologies can enable real-time engagement  
14  
15 323 through crowdsourcing<sup>43,44</sup> for ethical active living surveillance.<sup>8</sup>  
16  
17  
18  
19

## 20 324 **Strengths and Limitations**

21  
22  
23 325 The primary strength of the study is the development of novel and replicable methodology to  
24  
25 326 capture prospective PA from large populations using citizen-owned devices. This citizen science  
26  
27 327 approach, if replicated appropriately, can transform surveillance of physical PA among large  
28  
29 328 populations by leveraging citizen owned-devices. Implementing such innovative approaches of  
30  
31 329 PA surveillance will be critical to develop appropriate interventions to address global physical  
32  
33 330 inactivity.  
34  
35  
36

37 331 In terms of limitations, all observations are self-reported by citizen scientists. The study  
38  
39 332 sample size was also small after applying the inclusion criteria, however, smaller sample sizes  
40  
41 333 are not uncommon in smartphone-based EMA studies.<sup>45</sup> Another limitation is that IPAQ and  
42  
43 334 EMAs measured PA in different timeframes. As IPAQ captures data retrospectively and EMAs  
44  
45 335 capture data prospectively, they cannot be issued simultaneously. Nonetheless, although IPAQ  
46  
47 336 could have been issued on day 8, we refrained from such late deployment based on the evidence  
48  
49  
50  
51  
52  
53

337 from our pilots, which showed that compliance to burdensome traditional recall surveys such as  
338 IPAQ is much higher when it is issued as close to participant enrolment in the study as possible.

## 339 CONCLUSION

340 With growth of smartphones projected to only magnify in the future<sup>16</sup>, these ubiquitous tools  
341 can be leveraged via citizen science to capture accurate active living patterns of large populations  
342 in free-living conditions through innovative EMAs. This digital epidemiological and citizen  
343 science methodology adapted mobile EMAs to minimize recall bias and capture not only  
344 prospective PA, but also important physical and social contexts within which individuals  
345 accumulate PA.

## 346 LIST OF ABBREVIATIONS

347 EMAs: Ecological momentary assessments

348 PA: Physical activity

349 IPAQ: International physical activity questionnaire

## 350 DECLARATIONS

351 **Ethics approval and consent to participate:** All citizen scientists provided informed consent  
352 through the app and confirmed their age ( $\geq 18$  years) before being recruited. Ethics approval was  
353 obtained from the universities of Regina and Saskatchewan through a synchronized review  
354 protocol (REB # 2017-29).

355 **Consent for publication:** Not applicable

1  
2  
3 356 **Availability of data and materials:** The corresponding author will make the data available upon  
4  
5 357 reasonable request.  
6  
7

8 358 **Competing interests:** Authors declare no conflict of interest.  
9

10  
11 359 **Funding information:** This study was funded by the Saskatchewan Health Research  
12  
13 360 Foundation's Establishment Grant. This funding body had no role in study design; data  
14  
15 361 collection, analysis and interpretation; or in writing the manuscript.  
16

17 362 **Authors' contributions:** TRK contributed substantially to the study design, acquisition and  
18  
19 363 interpretation of data, and writing the manuscript. LMC contributed substantially to the  
20  
21 364 acquisition and interpretation of data, and writing the manuscript. All authors read and approved  
22  
23 365 the final manuscript.  
24

25 366 **Acknowledgments:** The authors acknowledge the entire DEPtH Lab team, and especially Ms  
26  
27 367 Kayla Brodersen for her role in the successful implementation of the SMART Platform.  
28  
29

30 368  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 369 REFERENCES

- 370 1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin*  
371 *Psychol.* 2008;4:1-32.
- 372 2. Burke LE, Shiffman S, Music E, et al. Ecological momentary assessment in behavioral  
373 research: addressing technological and human participant challenges. *J Med Internet Res.*  
374 2017;19(3):e77.
- 375 3. Thomas JG, Bond DS, Ryder BA, et al. Ecological momentary assessment of  
376 recommended postoperative eating and activity behaviors. *Surg Obes Relat Dis.* Mar-Apr  
377 2011;7(2):206-212.
- 378 4. Knell G, Gabriel KP. Ecological Momentary Assessment of Physical Activity: Validation  
379 Study. Jul 18 2017;19(7):e253.
- 380 5. Dunton GF. Ecological Momentary Assessment in Physical Activity Research. *Exerc*  
381 *Sport Sci Rev.* Jan 2017;45(1):48-54.
- 382 6. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A  
383 comparison of direct versus self-report measures for assessing physical activity in adults:  
384 a systematic review. *Int J Behav Nutr Phys Act.* Nov 6 2008;5:56.
- 385 7. Marszalek J, Morgulec-Adamowicz N, Rutkowska I, Kosmol A. Using Ecological  
386 Momentary Assessment to Evaluate Current Physical Activity. *BioMed Res.*  
387 2014;2014:9.
- 388 8. Katapally TR, Bhawra J, Leatherdale ST, et al. The SMART study, a mobile health and  
389 citizen science methodological platform for active living surveillance, integrated  
390 knowledge translation, and policy interventions: Longitudinal Study. *JMIR Public Health*  
391 *& Surveill.* 2018;4(1):e31.
- 392 9. Kanning M, Schlicht W. Be active and become happy: an ecological momentary  
393 assessment of physical activity and mood. *J Sport Exerc Psychol.* Apr 2010;32(2):253-  
394 261.
- 395 10. Conroy DE, Elavsky S, Hyde AL, Doerksen SE. The dynamic nature of physical activity  
396 intentions: a within-person perspective on intention-behavior coupling. *J Sport Exerc*  
397 *Psychol.* Dec 2011;33(6):807-827.
- 398 11. Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving self-reports of  
399 active and sedentary behaviors in large epidemiologic studies. *Exerc Sport Sci Rev.* Jul  
400 2012;40(3):118-126.
- 401 12. Adams SA, Matthews CE, Ebbeling CB, et al. The effect of social desirability and social  
402 approval on self-reports of physical activity. *Am J Epidemiol.* Feb 15 2005;161(4):389-  
403 398.
- 404 13. Hekler EB, Buman MP. Validation of Physical Activity Tracking via Android  
405 Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living  
406 Validation Studies. Apr 15 2015;3(2):e36.
- 407 14. Rodriguez VH, Medrano C, Plaza I, Corella C, Abarca A, Julian JA. Estimating the  
408 Physical Activity with Smartphones: Analysis of the Device Position and Comparison  
409 with GT3X+ Actigraph. 2016; Cham.

- 1  
2  
3 410 15. Silvertown J. A new dawn for citizen science. *Trends in ecology & evolution*.  
4 411 2009;24(9):467-471.
- 5 412 16. Bejarano CM, Cushing CC, Crick CJ. Does context predict psychological states and  
6 413 activity? An ecological momentary assessment pilot study of adolescents. *Psychol Sport*  
7 414 *Exerc*. 2019/03/01/ 2019;41:146-152.
- 8 415 17. Katapally TR. The SMART Framework: Integration of Citizen Science, Community-  
9 416 Based Participatory Research, and Systems Science for Population Health Science in the  
10 417 Digital Age. *JMIR Mhealth Uhealth*. Aug 30 2019;7(8):e14056.
- 11 418 18. T.R. K. SMART: A Mobile Health and Citizen Science Platform. [Internet]. 2019;  
12 419 <https://www.smartstudysask.com/copy-of-smart-youth>. Accessed November 22, 2019.
- 13 420 19. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire:  
14 421 12-country reliability and validity. *Med Sci Sports Exerc*. Aug 2003;35(8):1381-1395.
- 15 422 20. Sullivan R, Kinra S, Ekelund U, et al. Evaluation of the Indian Migration Study Physical  
16 423 Activity Questionnaire (IMS-PAQ): a cross-sectional study. *Int J Behav Nutr Phys Act*.  
17 424 Feb 9 2012;9:13.
- 18 425 21. Marszalek J, Morgulec-Adamowicz N, Rutkowska I, Kosmol A. Using ecological  
19 426 momentary assessment to evaluate current physical activity. *Biomed Res Int*.  
20 427 2014;2014:915172.
- 21 428 22. Bedard C, King-Dowling S. Understanding Environmental and Contextual Influences of  
22 429 Physical Activity During First-Year University: The Feasibility of Using Ecological  
23 430 Momentary Assessment in the MovingU Study. May 31 2017;3(2):e32.
- 24 431 23. Swendeman D, Comulada WS, Koussa M, et al. Longitudinal Validity and Reliability of  
25 432 Brief Smartphone Self-Monitoring of Diet, Stress, and Physical Activity in a Diverse  
26 433 Sample of Mothers. *JMIR mHealth and uHealth*. 2018;6(9):e176.
- 27 434 24. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity  
28 435 in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181.
- 29 436 25. Atienza AA, Moser RP, Perna F, et al. Self-reported and objectively measured activity  
30 437 related to biomarkers using NHANES. *Med Sci Sports Exerc*. 2011;43(5):815-821.
- 31 438 26. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a  
32 439 method of the future in clinical psychopharmacology. *Journal of Psychiatry and*  
33 440 *Neuroscience*. 2006;31(1):13.
- 34 441 27. Aadland E, Andersen LB, Anderssen SA, Resaland GK. A comparison of 10  
35 442 accelerometer non-wear time criteria and logbooks in children. *BMC Public Health*.  
36 443 2018/03/06 2018;18(1):323.
- 37 444 28. Katapally TR, Muhajarine N. Towards uniform accelerometry analysis: a standardization  
38 445 methodology to minimize measurement bias due to systematic accelerometer wear-time  
39 446 variation. *Journal of sports science & medicine*. 2014;13(2):379-386.
- 40 447 29. Liao Y, Intille SS, Dunton GF. Using ecological momentary assessment to understand  
41 448 where and with whom adults' physical and sedentary activity occur. *Int J Behav Med*. Feb  
42 449 2015;22(1):51-61.
- 43 450 30. Dollman J. Social and Environmental Influences on Physical Activity Behaviours.  
44 451 Multidisciplinary Digital Publishing Institute; 2018.

- 1  
2  
3 452 31. Benton JS, Anderson J, Hunter RF, French DP. The effect of changing the built  
4 453 environment on physical activity: a quantitative review of the risk of bias in natural  
5 454 experiments. *Int. J. Behav. Nutr. Phys. Act.* 2016;13(1):107.
- 6 455 32. Coghill C-L, Valaitis RK, Eyles JD. Built environment interventions aimed at improving  
7 456 physical activity levels in rural Ontario health units: a descriptive qualitative study. *BMC*  
8 457 *public health.* 2015;15(1):464.
- 9 458 33. Wilkie S, Townshend T, Thompson E, Ling J. Restructuring the built environment to  
10 459 change adult health behaviors: a scoping review integrated with behavior change  
11 460 frameworks. *Cities & health.* 2018;2(2):198-211.
- 12 461 34. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile  
13 462 technology into psychosocial and health behaviour treatments. *British journal of health*  
14 463 *psychology.* 2010;15(1):1-39.
- 15 464 35. Daugherty DA, Runyan JD, Steenbergh TA, Fratzke BJ, Fry BN, Westra E. Smartphone  
16 465 delivery of a hope intervention: Another way to flourish. *PloS one.* 2018;13(6):e0197930.
- 17 466 36. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAs)  
18 467 in mobile health: key components and design principles for ongoing health behavior  
19 468 support. *Ann. Behav. Med.* 2017;52(6):446-462.
- 20 469 37. Zink J, Belcher BR, Dzibur E, et al. Association between self-reported and objective  
21 470 activity levels by demographic factors: ecological momentary assessment study in  
22 471 children. *JMIR mHealth and uHealth.* 2018;6(6):e150.
- 23 472 38. Dunton GF, Liao Y, Intille SS, Spruijt-Metz D, Pentz M. Investigating children's physical  
24 473 activity and sedentary behavior using ecological momentary assessment with mobile  
25 474 phones. *Obesity.* 2011;19(6):1205-1212.
- 26 475 39. Bonney R, Cooper CB, Dickinson J, et al. Citizen science: a developing tool for  
27 476 expanding science knowledge and scientific literacy. *BioScience.* 2009;59(11):977-984.
- 28 477 40. Statista. Smartphone users worldwide 2020 Statista. [Internet]. 2019;  
29 478 <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/>.  
30 479 Accessed January 16, 2020.
- 31 480 41. Mittelstadt B, Benzler J, Engelmann L, Prainsack B, Vayena E. Is there a duty to  
32 481 participate in digital epidemiology? *Life Sciences, Society and Policy.* 2018/05/09  
33 482 2018;14(1):9.
- 34 483 42. Hinckson E, Schneider M, Winter SJ, et al. Citizen science applied to building healthier  
35 484 community environments: advancing the field through shared construct and measurement  
36 485 development. *Int. J. Behav. Nutr. Phys. Act.* 2017;14(1):133.
- 37 486 43. Kraemer JD, Strasser AA, Lindblom EN, Niaura RS, Mays D. Crowdsourced data  
38 487 collection for public health: A comparison with nationally representative, population  
39 488 tobacco use data. *Preventive medicine.* 2017;102:93-99.
- 40 489 44. Behrend TS, Sharek DJ, Meade AW, Wiebe EN. The viability of crowdsourcing for  
41 490 survey research. *Behavior research methods.* 2011;43(3):800.
- 42 491 45. Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of  
43 492 direct versus self-report measures for assessing physical activity in adults: a systematic  
44 493 review. *Int. J. Behav. Nutr. Phys. Act.* 2008;5(1):56.
- 45 494

495 **Table 1. Demographic characteristics of the final sample**

<b>Demographic Characteristics</b>	<b>Categories</b>	<b>n</b>	<b>%</b>
Sex	Male	26	29.21
	Female	46	51.68
	Did not identify	17	19.11
Age in years, mean (SD)		71	37.15 (15.92)
Body Mass Index, mean (SD)		73	28.46 (7.78)
Annual Household income	< 40,000	14	15.73
	40,000- < 70,000	21	23.60
	≥ 70,000	35	39.32
	Did not respond	19	21.35
Educational attainment	Some or completed secondary/high school	7	7.86
	Some post-secondary (university or college)	18	20.22
	Received university or college degree/diploma	46	51.68
	Did not respond	18	20.24

496 Note: SD: Standard deviation

497

498

499



500 **Table 2. Overall physical activity and intensity measurements: IPAQ vs. EMA**

	Mean (minutes/day)	Standard Deviation	Percentiles (minutes/day)			p- value*
			25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	
<b>Overall physical activity measurement: IPAQ vs. EMA</b>						
IPAQ	194.39	266.10	63.80	122.14	175.72	0.331
EMA	140.91	98.31	73.07	123.75	183.48	
<b>Light and moderate-to-vigorous physical activity measurement: IPAQ vs. EMA</b>						
Light PA (IPAQ)	116.99	171.24	36.00	67.86	110.00	0.322
Light PA (EMA)	87.16	64.44	41.25	68.33	103.67	
Moderate to vigorous (IPAQ)	98.42	175.18	17.14	49.44	92.86	0.995
Moderate to vigorous (EMA)	70.38	63.48	40.00	52.50	87.50	

501 Note: Based on Wilcoxon Signed Ranks Test; IPAQ: international physical activity  
 502 questionnaire; EMA: ecological momentary assessment  
 503

504 **Table 3. Spearman correlation coefficients between IPAQ and EMA across physical**  
 505 **activity intensities**

Intensity	Spearman correlation coefficients	
	$\rho$ (p-value)	n
Overall PA	0.414 (0.001)	89
Light	0.258 (0.012)	87
Moderate-to-vigorous	0.316 (0.009)	67



512 **Table 4. Overall physical activity measurement across physical domains: IPAQ vs. EMA**

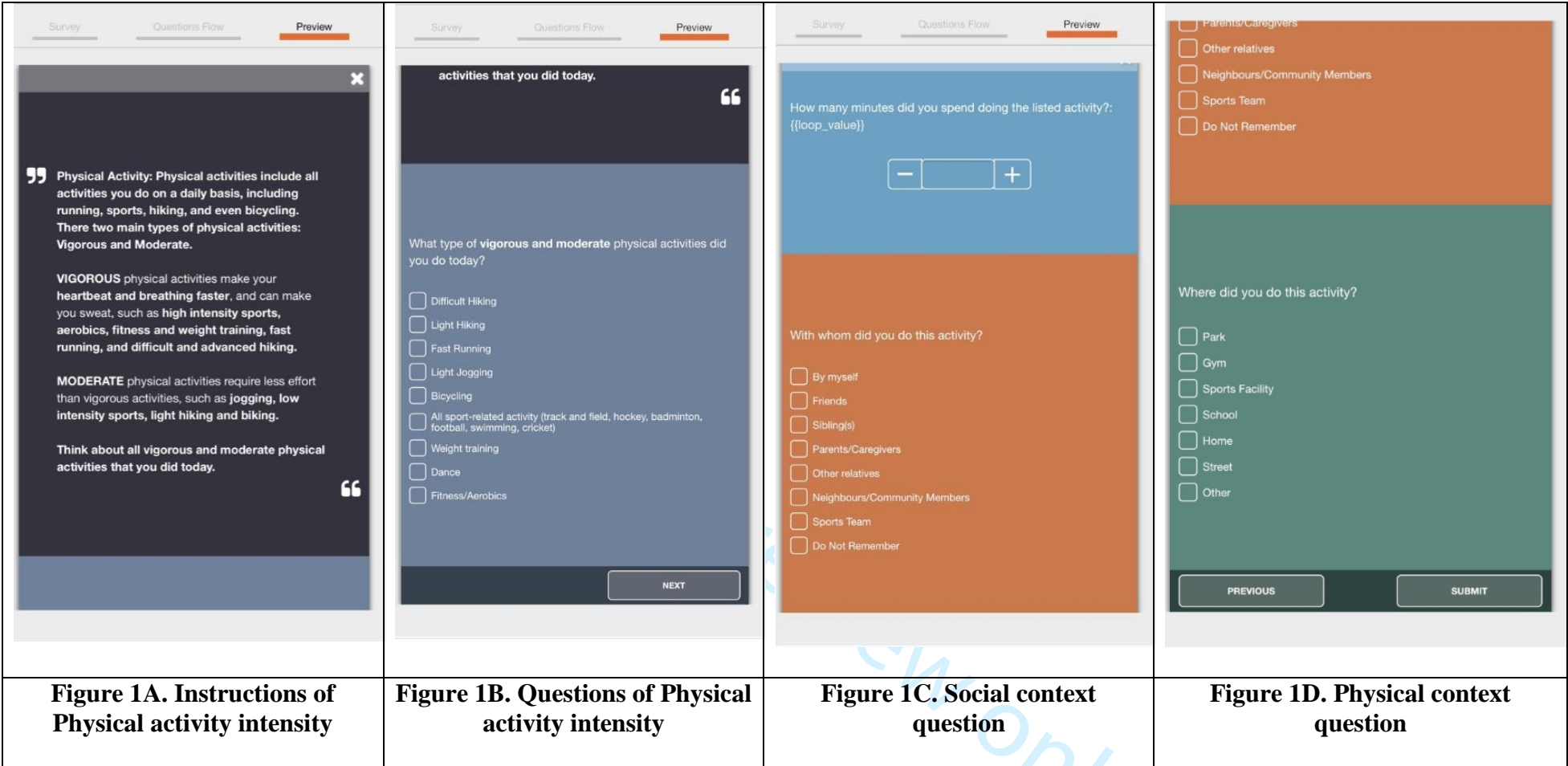
Domain		Mean	SD	Percentiles (minutes/day)			p-value*
		(minutes/day)		25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	
Workplace	Survey	32.14	34.97	7.86	12.86	66.07	0.345
	EMA	20.50	17.87	5.63	15.00	38.13	
Active Transportation PA	Survey	43.97	24.32	25.36	40.00	66.43	0.002
	EMA	16.41	11.51	7.56	10.00	25.69	
Household	Survey	38.27	35.01	9.04	28.50	74.46	0.117
	EMA	25.33	46.29	5.16	10.31	19.84	
Recreation Sport and Leisure-time PA	Survey	145.90	306.95	15.00	34.29	72.86	0.003
	EMA	20.88	16.19	7.50	17.50	37.50	

513 Note: Based on related-samples Wilcoxon Signed Rank Tests; IPAQ: international physical  
 514 activity questionnaire; EMA: ecological momentary assessment

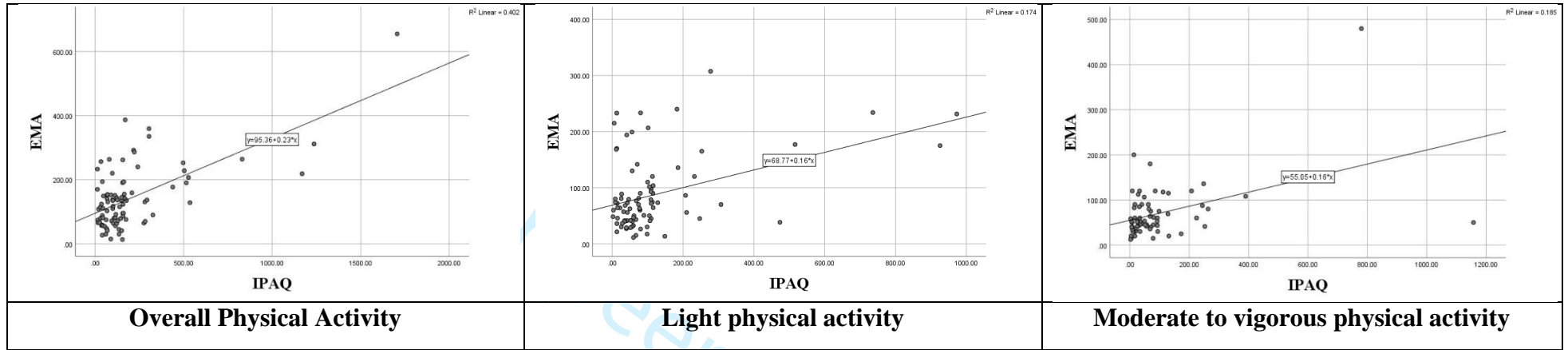
515  
 516  
 517 **Table 5. Spearman correlation coefficients of overall physical activity between EMA and**  
 518 **IPAQ across physical domains**

Domain	Spearman correlation coefficients	
	$\rho$ (p-value)	n
Physical Activity (PA) at workplace	0.500 (0.391)	5
Transportation PA	0.166 (0.587)	13
Housework, house maintenance and caring for family	0.607 (0.036)	12
Recreation, Sport and Leisure-time PA	0.587 (0.021)	15

519



**Figure 1. Time-triggered ecological momentary assessment capturing prospective physical activity**



**Figure 2. Correlation between IPAQ and EMA measurements of physical activity**

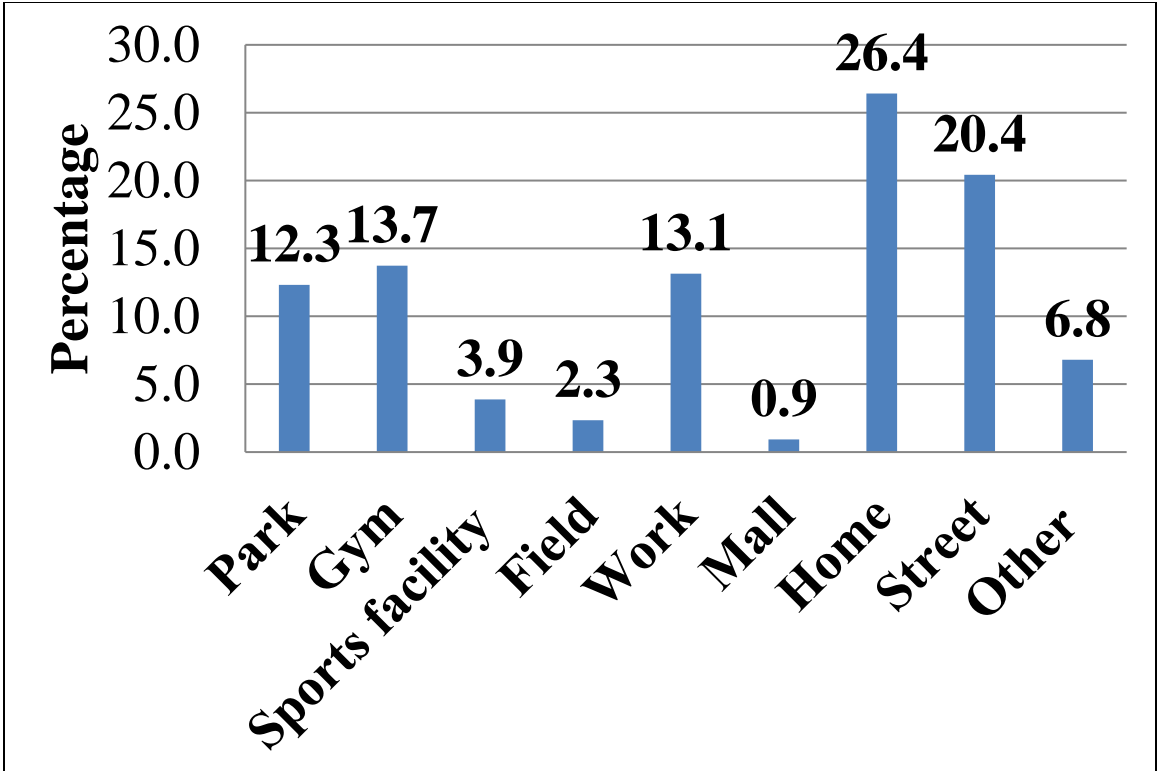


Figure 3. Distribution of Daily EMA Physical activity within physical contexts

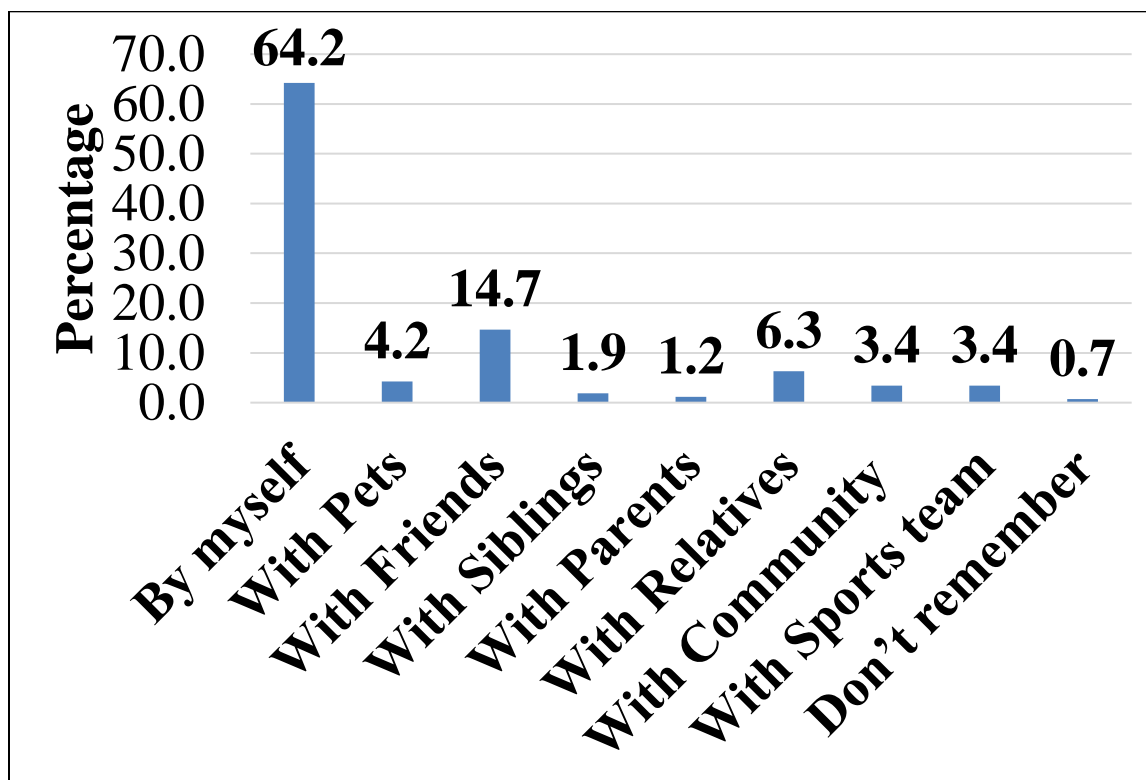


Figure 4. Distribution of Daily EMA Physical activity within social contexts

## STROBE Statement

Checklist of items that should be included in reports of observational studies

Section/Topic	Item No	Recommendation	Reported on Page No
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
<b>Introduction</b>			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
<b>Methods</b>			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6
		<i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	
		<i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	
Variables	7	(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed	6,7,8,9
		<i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Data sources/measurement	8*	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6,7,8,9
Bias	9	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6,7
Study size	10	Describe any efforts to address potential sources of bias	9
Quantitative variables	11	Explain how the study size was arrived at	9
Statistical methods	12	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6,7,8,9
		(a) Describe all statistical methods, including those used to control for confounding	9
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	9
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed	
		<i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed	9
		<i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	

Section/Topic	Item No	Recommendation	Reported on Page No
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	10, 11
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10,11
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key results	18	Summarise key results with reference to study objectives	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12.13
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
<b>Other Information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	17

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).