Using smartphones and wearable devices to monitor behavioral changes during COVID-19

Shaoxiong Sun, Amos A Folarin, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Nicholas Cummins, Faith Matcham, Gloria Dalla Costa, Sara Simblett, Letizia Leocani, Femke Lamers, Per Soelberg Sørensen, Mathias Buron, Ana Zabalza, Ana Isabel Guerrero Pérez, Brenda WJH Penninx, Sara Siddi, Josep Maria Haro, Inez Myin-Germeys, Aki Rintala, Til Wykes, Vaibhav A. Narayan, Giancarlo Comi, Matthew Hotopf, Richard J. B. Dobson

Original Paper

Using Smartphones and Wearable Devices to Monitor Behavioral Changes During COVID-19

Shaoxiong Sun[1], PhD; Amos A Folarin[1,2], PhD; Yatharth Ranjan[1], MSc; Zulqarnain Rashid[1], PhD; Pauline Conde[1], BSc; Callum Stewart[1], MSc; Nicholas Cummins[3], PhD; Faith Matcham[4], PhD; Gloria Dalla Costa[5], MD; Sara Simblett[6], PhD; Letizia Leocani[5], MD, PhD; Femke Lamers[7], PhD; Per Soelberg Sørensen[8], MD; Mathias Buron[8], MD; Ana Zabalza[9], MD; Ana Isabel Guerrero Pérez[9], MSc; Brenda WJH Penninx[7], PhD; Sara Siddi[10,11,12], PhD; Josep Maria Haro[10,11,12], MD; Inez Myin-Germeys[13], PhD; Aki Rintala[13], MSc; Til Wykes[6,14], PhD; Vaibhav A Narayan[15], PhD; Giancarlo Comi[16], MD; Matthew Hotopf[4,14], PhD; Richard JB Dobson[1,2], PhD; RADAR-CNS Consortium[17]

[1]The Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[2]Institute of Health Informatics, University College London, London, United Kingdom

[3]Chair of Embedded Intelligence for Health Care & Wellbeing, University of Augsburg, Augsburg, Germany

[4]The Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[5]Neurorehabilitation Unit and Institute of Experimental Neurology, University Vita Salute San Raffaele, Istituto Di Ricovero e Cura a Carattere Scientifico Ospedale San Raffaele, Milan, Italy

[6]The Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[7]Department of Psychiatry, Amsterdam Public Health Research Institute and Amsterdam Neuroscience, Amsterdam University Medical Centre, Vrije Universiteit and GGZ inGeest, Amsterdam, Netherlands

[8]Danish Multiple Sclerosis Centre, Department of Neurology, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark

[9]Multiple Sclerosis Centre of Catalonia (Cemcat), Department of Neurology/Neuroimmunology, Hospital Universitari Vall d'Hebron, Barcelona, Spain

[10]Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, Barcelona, Spain

[11]Centro de Investigación Biomédica en Red de Salud Mental, Madrid, Spain

[12]Universitat de Barcelona, Barcelona, Spain

[13]Centre for Contextual Psychiatry, Department of Neurosciences, Katholieke Universiteit Leuven, Leuven, Belgium

[14]South London and Maudsley National Health Services Foundation Trust, London, United Kingdom

[15]Janssen Research and Development LLC, Titusville, NJ, United States

[16]Institute of Experimental Neurology, Istituto Di Ricovero e Cura a Carattere Scientifico Ospedale San Raffaele, Milan, Italy

[17]The RADAR-CNS Consortium, London, United Kingdom

Corresponding Author:
Shaoxiong Sun, PhD
The Department of Biostatistics and Health Informatics
Institute of Psychiatry, Psychology and Neuroscience
King's College London
Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London
PO Box 80 De Crespigny Park, Denmark Hill
London, SE58AF
United Kingdom
Phone: 44 02078480951
Email: shaoxiong.sun@kcl.ac.uk

Abstract

Background: In the absence of a vaccine or effective treatment for COVID-19, countries have adopted nonpharmaceutical interventions (NPIs) such as social distancing and full lockdown. An objective and quantitative means of passively monitoring the impact and response of these interventions at a local level is needed.

Objective: We aim to explore the utility of the recently developed open-source mobile health platform Remote Assessment of Disease and Relapse (RADAR)–base as a toolbox to rapidly test the effect and response to NPIs intended to limit the spread of COVID-19.

Methods: We analyzed data extracted from smartphone and wearable devices, and managed by the RADAR-base from 1062 participants recruited in Italy, Spain, Denmark, the United Kingdom, and the Netherlands. We derived nine features on a daily basis including time spent at home, maximum distance travelled from home, the maximum number of Bluetooth-enabled nearby devices (as a proxy for physical distancing), step count, average heart rate, sleep duration, bedtime, phone unlock duration, and social app use duration. We performed Kruskal-Wallis tests followed by post hoc Dunn tests to assess differences in these features among baseline, prelockdown, and during lockdown periods. We also studied behavioral differences by age, gender, BMI, and educational background.

Results: We were able to quantify expected changes in time spent at home, distance travelled, and the number of nearby Bluetooth-enabled devices between prelockdown and during lockdown periods ($P<.001$ for all five countries). We saw reduced sociality as measured through mobility features and increased virtual sociality through phone use. People were more active on their phones ($P<.001$ for Italy, Spain, and the United Kingdom), spending more time using social media apps ($P<.001$ for Italy, Spain, the United Kingdom, and the Netherlands), particularly around major news events. Furthermore, participants had a lower heart rate ($P<.001$ for Italy and Spain; $P=.02$ for Denmark), went to bed later ($P<.001$ for Italy, Spain, the United Kingdom, and the Netherlands), and slept more ($P<.001$ for Italy, Spain, and the United Kingdom). We also found that young people had longer homestay than older people during the lockdown and fewer daily steps. Although there was no significant difference between the high and low BMI groups in time spent at home, the low BMI group walked more.

Conclusions: RADAR-base, a freely deployable data collection platform leveraging data from wearables and mobile technologies, can be used to rapidly quantify and provide a holistic view of behavioral changes in response to public health interventions as a result of infectious outbreaks such as COVID-19. RADAR-base may be a viable approach to implementing an early warning system for passively assessing the local compliance to interventions in epidemics and pandemics, and could help countries ease out of lockdown.

Introduction

On March 11, 2020, the World Health Organization declared the rapidly spreading SARS-CoV-2 virus outbreak a pandemic. This novel coronavirus is the cause of a contagious acute respiratory disease (COVID-19), which was first reported in Wuhan, Hubei Province, China [1-3]. As of July 1, 2020, it had infected over 10 million people and spread to 213 countries and territories around the world [4]. Although precise statistics on mortality are being determined, COVID-19 can be deadly with an estimated 1% case fatality rate, and this rate increases for older adults and those with underlying health problems [5,6]. The outbreak of COVID-19 has placed an unprecedented burden on health care systems in most-affected countries and has resulted in considerable economic losses and a possible global recession [7,8].

To date, there is no vaccine or highly effective treatment. The widely adopted strategy has been the use of nonpharmaceutical interventions (NPIs) such as social distancing and even full lockdown to control the spread of the virus and ease the pressure on health and care systems [9,10]. NPIs have been implemented in many countries including China, Italy, Spain, the United Kingdom, and the Netherlands. These measures have been shown to considerably reduce the new confirmed cases [9]. Key to the success of NPIs is the timing of these interventions and the response of the population, both of which might differ among countries and could necessitate further interventions in the case of low compliance either nationally or locally. Furthermore,

US $11 trillion of fiscal measures have been announced by more than two-thirds of governments worldwide in an attempt to mitigate the fallout from the pandemic and lockdown [11]. Therefore, we urgently require an objective and quantitative way to monitor population behavior to assess the impact and response of such interventions. Additionally, we need to monitor for the potential effects of a rebound in cases in the winter months as social distancing measures are relaxed and to strategize and understand where course corrections are required. Similarly, understanding potential seasonal forcing of COVID-19 will require a good understanding of different NPIs' effects, so they can be factored out.

The increasing availability of wide-bandwidth mobile networks, smartphones, and wearable sensors makes it possible to collect near real-time high-resolution data sets from large numbers of participants and greatly facilitates remote monitoring of behavior [12-14]. By leveraging sensor modalities in smartphones, which includes network and GPS location tracking, and Fitbit devices, which includes step counts and heart rate, it is possible to access mobility and even wellness for the population. To manage the data collected from multiple sensor modalities and mobile devices, platforms such as the open-source Remote Assessment of Disease and Relapse (RADAR)–base [15] mobile health platform have been developed [16]. This platform has been used to enable remote monitoring in a range of use cases including central nervous system diseases (major depressive disorder [MDD], epilepsy, and multiple sclerosis [MS]) as part of the

Innovative Medicines Initiative (IMI2) RADAR–Central Nervous System (CNS) major program [17,18].

In this paper, we explore the utility of the RADAR-base platform as a toolbox to test the effect and response of NPIs aimed at limiting the spread of infectious diseases such as COVID-19 by leveraging participant data already collected from November 2017 onward as part of the ongoing RADAR-CNS studies [16,17,19]. Specifically, we created measures of mobility (as a proxy of physical distancing), phone use (as a proxy of virtual sociality), and physiological measures (heart rate and sleep), and compared these features among the baseline, prelockdown, and during lockdown periods. Furthermore, we also provide a joint analysis of these features to provide a holistic view and interpret these behavioral changes during COVID-19.

Methods

Data Collection

The RADAR-CNS studies were approved by all local ethics committees, and all participants signed informed consent [19]. We included 1062 participants recruited in five European countries: Italy, Spain, Denmark, the United Kingdom, and the Netherlands. Participants in the Netherlands were partially recruited through Hersenonderzoek.nl [20]. The data were collected for the purpose of finding new ways of monitoring MDD (Spain: n=150; the Netherlands: n=103; and the United Kingdom: n=316) and MS (Milan, Italy: n=208; Barcelona, Spain: n=179; and Copenhagen, Denmark: n=106) using wearable devices and smartphone technology to improve patients' quality of life (QoL) and potentially change the treatment of these and other chronic disorders. As we focused on country-level behavioral changes in response to the NPIs, we aggregated data collected in Spain and did not focus on analyzing differences between participants with MDD and MS. Passive participant data, that is data that did not require conscious participant engagement, were collected continuously on a 24/7 basis through a smartphone and a Fitbit device, which included location, Bluetooth, activity, sleep, heart rate, and phone use data. In this study, we used participants' own Android smartphones where available and provided a participant with a Motorola G5, G6, or G7 if participants had an iPhone or did not have a smartphone. For Fitbit devices, Fitbit Charge 2 devices were given to participants, and then Fitbit Charge 3 devices were given to the recently recruited participants when Fitbit Charge 2 devices were no longer available. We asked participants to wear the device on their nondominant hand. Although not used for this study, active data were also collected, which required clinicians or participants to fill out emailed surveys (eg, Inventory of Depressive Symptomatology [Self-Report]), app-delivered questionnaires (eg, Patient Health Questionnaire), or perform short clinical tests (eg, Expanded Disability Status Scale).

The data collection and management were handled by the open-source mHealth platform RADAR-Base [16]. The platform provides high scalability, interoperability, flexibility, and reliability while allowing the freedom for anyone to deploy. Due to the streaming first nature of the platform, it is also easy to aggregate, analyze, and provide insights into the data in real time, hence making the results of this work potentially deployable for localized monitoring and targeted interventions.

Feature Extraction

To study physical-behavioral changes in response to COVID-19 NPIs, we examined participants' mobility by analyzing relative location and Bluetooth data from smartphones and step count data from Fitbit devices. We investigated phone unlock duration and social app use duration to study online social-behavioral changes. Physiological measures such as sleep and heart rate from Fitbit devices were also analyzed to identify possible changes as a result of lockdown. A full list of features is presented in Table 1. These features were extracted for each participant every day. The daily features were calculated using the data from 6 AM on the present day to 6 AM on the next day for all features except total sleep duration and bedtime, where 8 PM was used as the starting time point and 11 AM the finishing. When no data were found in a data modality for a participant on a day due to the participant not wearing the Fitbit device or not using the smartphone, we did not calculate the feature derived from that data modality on that day.

Table 1. A full list of extracted features.

Category and modality	Features	Extraction
Mobility		
Smartphone location	• Maximum travelled distance from home • Homestay	• The maximum distance travelled from home location • The time spent within 200 m radius of home location (determined using DBSCAN[a])
Smartphone Bluetooth	• Maximum number of nearby devices	• The maximum number of Bluetooth-enabled nearby devices
Fitbit step count	• Step count	• Daily total of Fitbit step count
Physiological measures		
Fitbit sleep	• Bedtime • Sleep duration	• The first sleep category at night • Daily total duration of sleep categories (light, deep, and REM[b])
Fitbit heart rate	• Average heart rate	• The daily average heart rate
Phone use		
Smartphone user interaction	• Unlock duration	• The total duration of phone in the unlocked state
Smartphone use event	• Social app use duration	• The total duration spent on social apps (Google Play categories of Social, Communication, and Dating)

[a]DBSCAN: density-based spatial clustering of applications with noise.

[b]REM: rapid eye movement.

The smartphone-derived location data were sampled once every 5 minutes by default, with longer sampling durations dependent on network connectivity. Spurious location coordinates were identified and removed if they differed from preceding and following coordinates by more than five degrees. Home location was determined daily by clustering location data between 8 PM and 4 AM with the mean coordinate of the cluster that the last coordinate belonged to being used. This choice was made because the largest cluster may not be the home location for a single night but the last location before phones shut down had a higher probability to be home location for that night. The clustering was implemented using density-based spatial clustering of applications with noise [21]. A duration gated by two adjacent coordinates was regarded as a valid homestay duration on the condition that both coordinates were no further than 200 meters from the home location. A duration longer than 1 hour was excluded due to the large proportion of missing data when compared to the 5-minute sampling duration. All valid homestay durations between 8 AM and 11 PM were summed to calculate daily homestay. Daily maximum distance from home was also computed based on the coordinates in the same period.

Bluetooth data, including the number of nearby and paired devices, were also collected from smartphones, which were sampled every hour. The daily maximum number of nearby devices was used as a mobility feature. An increased number of nearby devices (typically other phones) detected may indicate other users' presence in the vicinity, which therefore can serve as a proxy of physical distancing.

In addition to mobility features extracted from smartphones, daily step count was taken from the Fitbit device, which was computed as the total steps a participant walked every day. Likewise, daily sleep duration was computed as the summation of three Fitbit-output sleep categories (light, deep, and rapid eye movement) sampled every 30 seconds from 8 PM to 11 AM the next day. Bedtime was defined as the time of the first sleep category reported by Fitbit after 8 PM. Note that the sleep categories referred to the sleep stages provided by the Fitbit application programming interface [22], which are not equivalent to the medical sleep stages. Finally, daily mean heart rate was calculated by averaging the Fitbit-output heart rate readings, sampled every 5 seconds at best. This sampling interval may be longer depending on Fitbit proprietary algorithms for remaining battery level, quality scoring, and network connectivity.

To explore changes in phone use, daily unlock duration was calculated by summing time periods starting with the unlocked state and ending with the standby state. Single intervals longer than 4 hours were excluded, which might result from a missing standby state or unintentionally leaving the phone unlocked. App use was quantified by classifying apps according to categories listed on Google Play [23]. As we were particularly interested in cyber-social interactions, we focused on the daily use time of social apps, including the Google Play categories of Social, Communication, and Dating. Among them are Facebook, Instagram, and WhatsApp.

Data Analysis

We plotted how the features evolved over 1.5 years for each country investigated. The participants' daily median, 25th percentile, and 75th percentile of each feature were calculated and then plotted. A minimum of 20 participants' data points was a prerequisite for calculation for any given day to reduce variance and noise. To facilitate interpretation, we also marked time points of public announcements related to lockdown policies [24].

To examine changes in mobility, physiological measures, and phone use induced by the lockdowns, comparisons among baseline, prelockdown, and during lockdown on the daily median of each feature were carried out using Kruskal-Wallis tests followed by post hoc Dunn tests [25,26]. For the during lockdown phase, we chose the entire period of the national lockdown in each country, which ended when NPIs were eased for the first time. For the prelockdown phase, we chose the period immediately prior to the first restrictive measure with the same length of the entire national lockdown. For the baseline phase, we chose the same period in 2019 as the 2020 national lockdown for countries starting to collect data earlier than 2019, which included Italy, Spain, and the United Kingdom. This was aimed at suppressing seasonal variability. For Denmark and the Netherlands where participant recruitment and data collection started much later, we chose the period that started with the earliest stable date (no considerate missing data or outliers) with the same length of the entire national lockdown. If a significant difference among these three periods was found after Benjamini–Yekutieli correction for the number of features (n=9), post hoc Dunn test was applied with Benjamini–Yekutieli correction for the number of groups (n=3) [27]. Box plots were used to present the results. A $P<.05$ after Benjamini–Yekutieli corrections was deemed statistically significant. It should be noted that we applied corrections resulting from multiple comparisons and multiple features in each country.

We also studied factors that might influence the subpopulation behavioral features during the lockdown period. The investigated factors included age, gender, BMI, and educational background. For age groups, we defined the young group as younger than 45 years and the older adult group as 45 years or older. For BMI groups, the low BMI group was defined as less than 25, and the high BMI group as greater than or equal to 25. For education groups, we defined the degree group as having a bachelor's degree or above and the nondegree group as having lower qualifications. Furthermore, we defined a combined factor group of young men, as this subpopulation was suspected to be less compliant with social distancing measures. Here we focused on features of homestay and daily step count during the entire period of lockdown for each country. We performed Wilcoxon signed rank tests on these two features to examine statistically significant differences. The P values were corrected with the number of factors (n=5) and the number of features (n=2) using Benjamini–Yekutieli correction.

Finally, we investigated the effects of different NPIs, in particular immediately after national lockdowns. This was done by comparing the NPIs implemented in the five countries within the first 2 weeks after entering national lockdowns.

Results

Plots showing how the extracted features evolved from February 1, 2019, to July 5, 2020, and box plots of these features are shown in Figures 1-5 and in Figure 6, respectively. Detailed test statistics and P values comparing prelockdown and during lockdown measures are presented in Table 2. Figure 7 shows a zoomed in version of Figures 3 and 4.

Figure 1. Behavioral changes for Milan, Italy (208 participants). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.
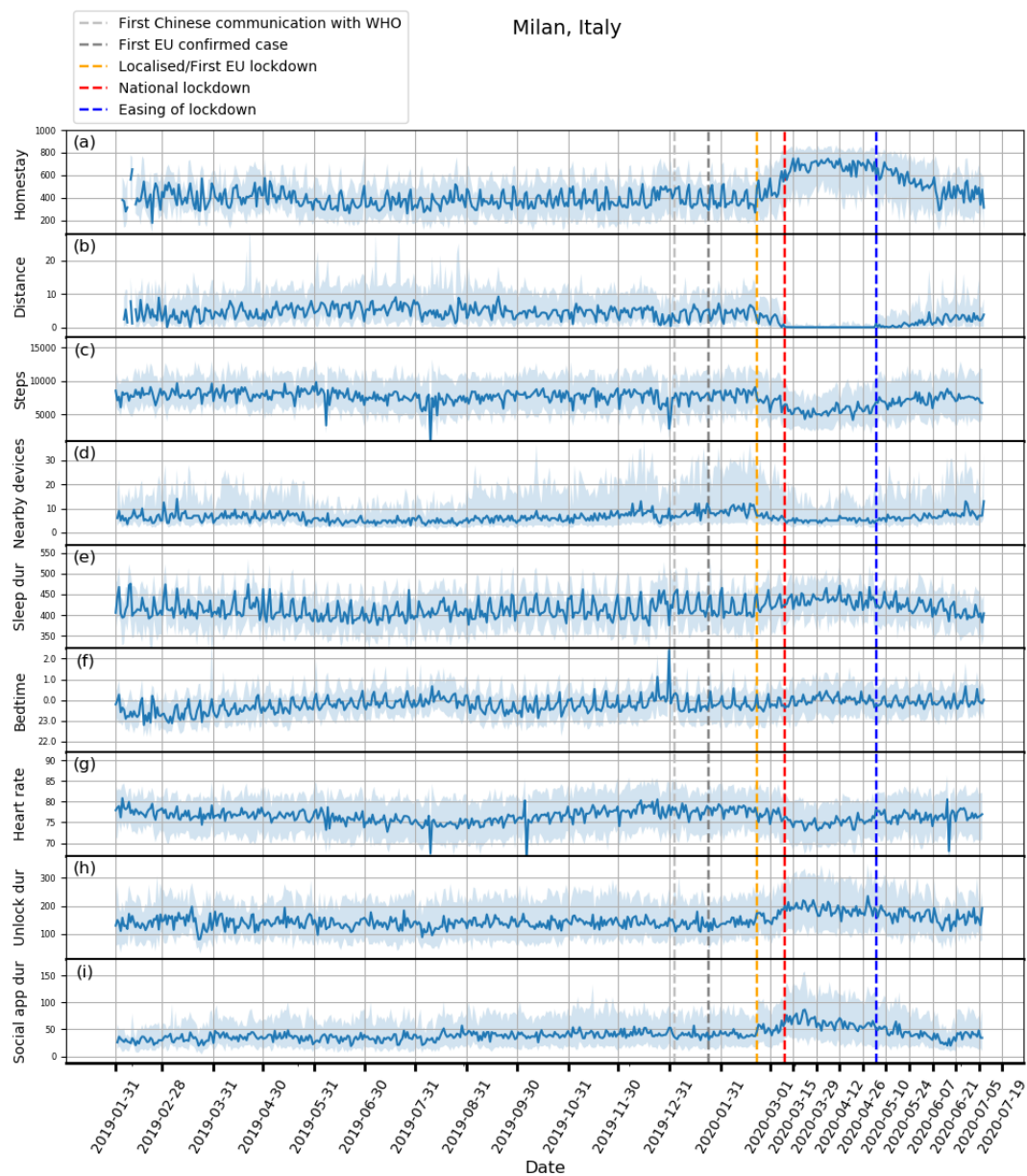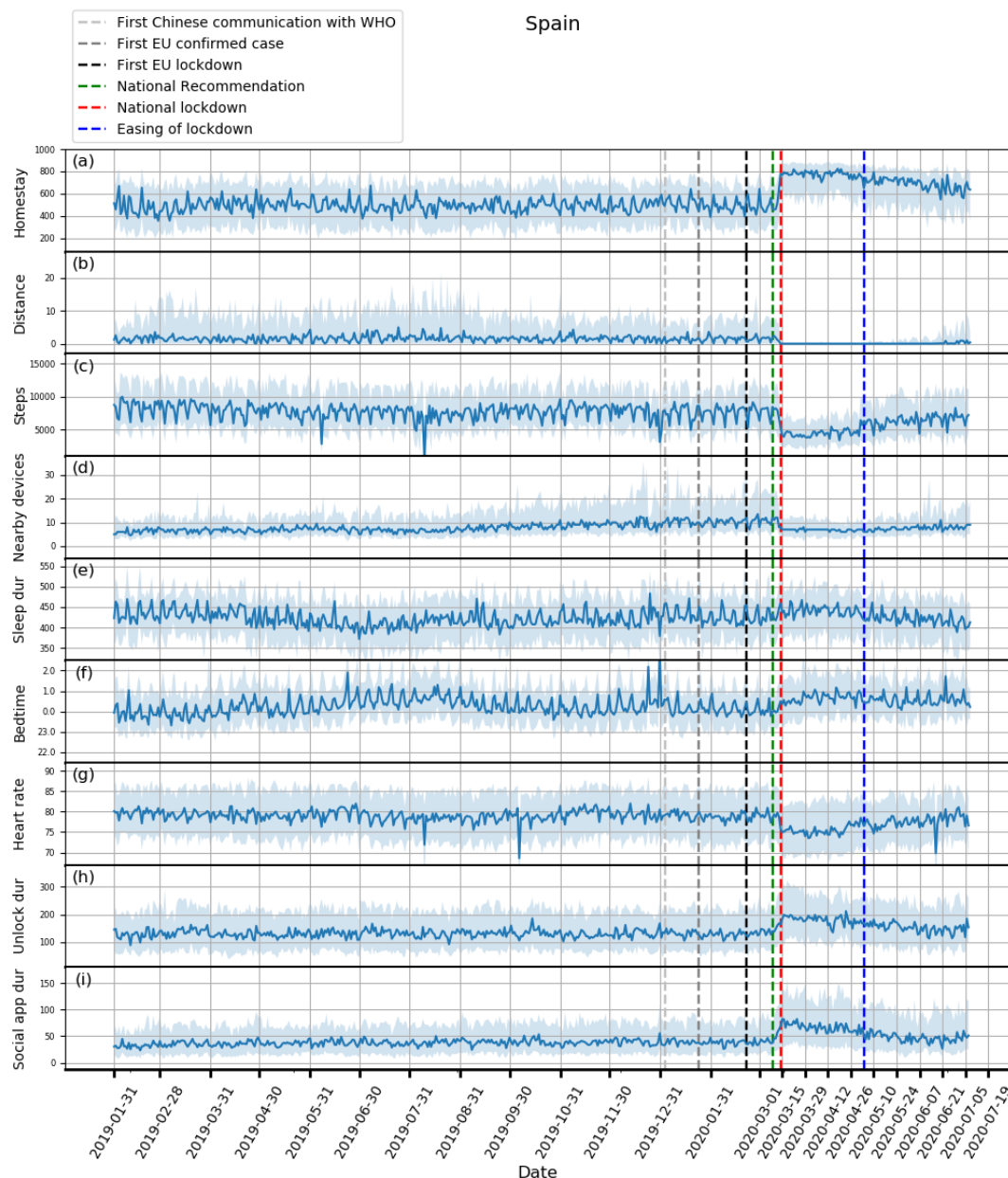
Figure 2. Behavioral changes for Spain (329 participants). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.


Through RADAR-base, we quantified changes in mobility, phone use, and physiological measures as a result of NPIs introduced to control COVID-19. As expected, following national lockdowns, participants in all countries stayed at home for longer, travelled shorter distances, walked less, and had fewer Bluetooth-enabled devices in the vicinity.

In contrast to increased physical distancing (reduced sociability) suggested by these mobility features, higher phone use, indicating compensatory sociability, was observed. Italy, Spain, and the United Kingdom saw longer unlock duration, and these 3 countries together with the Netherlands also showed longer

social app use duration. Tellingly, both unlock duration and social app use duration saw peaks around the news of national lockdowns in all countries.

Concurrent with the changes in mobility and phone use, changes in physiological measures were observed. Participants in Spain, Italy, and the United Kingdom went to bed later and slept more. Participants in Spain, Italy, and Denmark also had a decrease in heart rate. Although not statistically significant, an increase in sleep duration and bedtime in Denmark and the Netherlands, and a decrease in heart rate in the United Kingdom and the Netherlands can be seen in Figures 3-5.

Figure 3. Behavioral changes for Copenhagen, Denmark (106 participants). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.
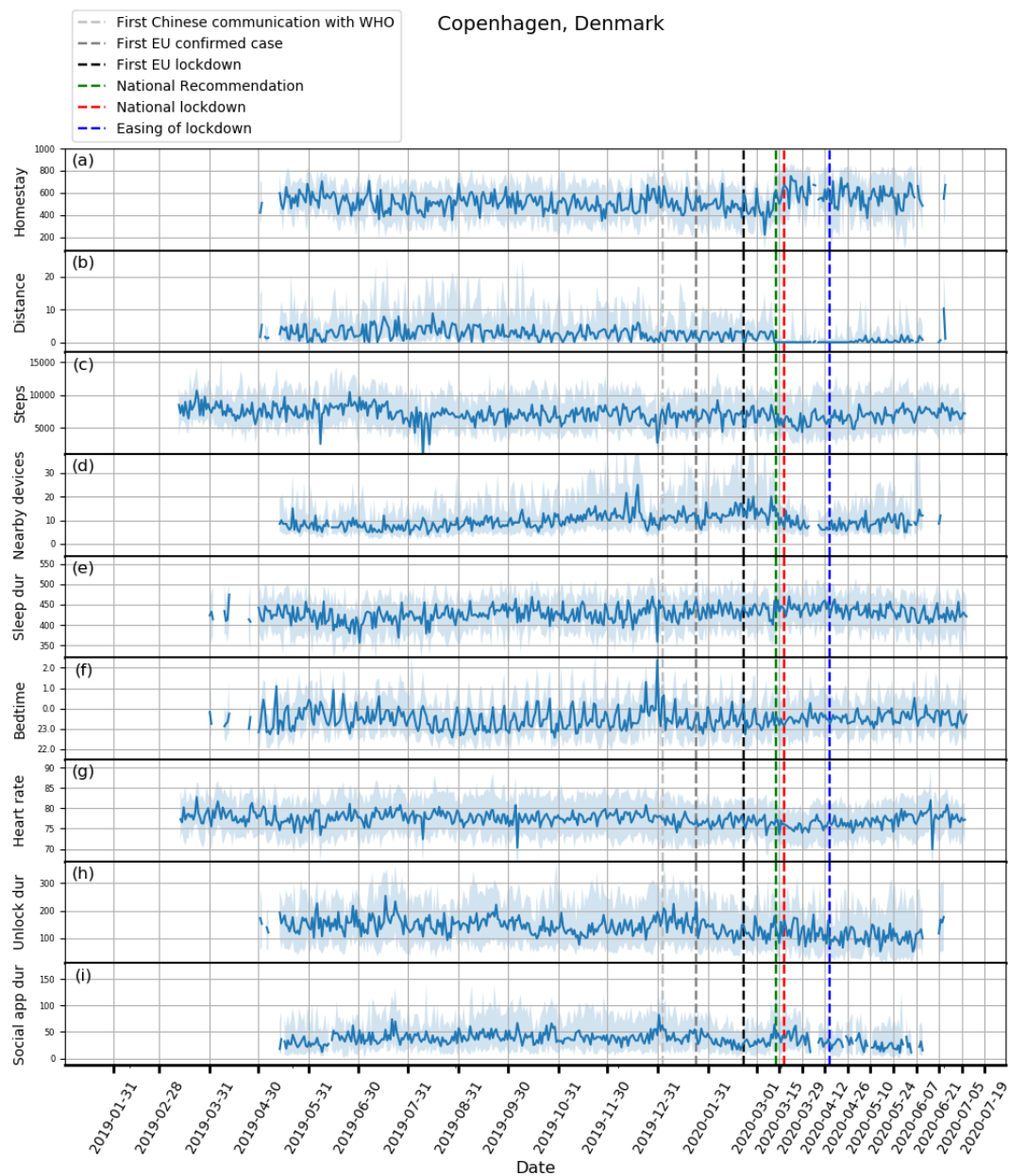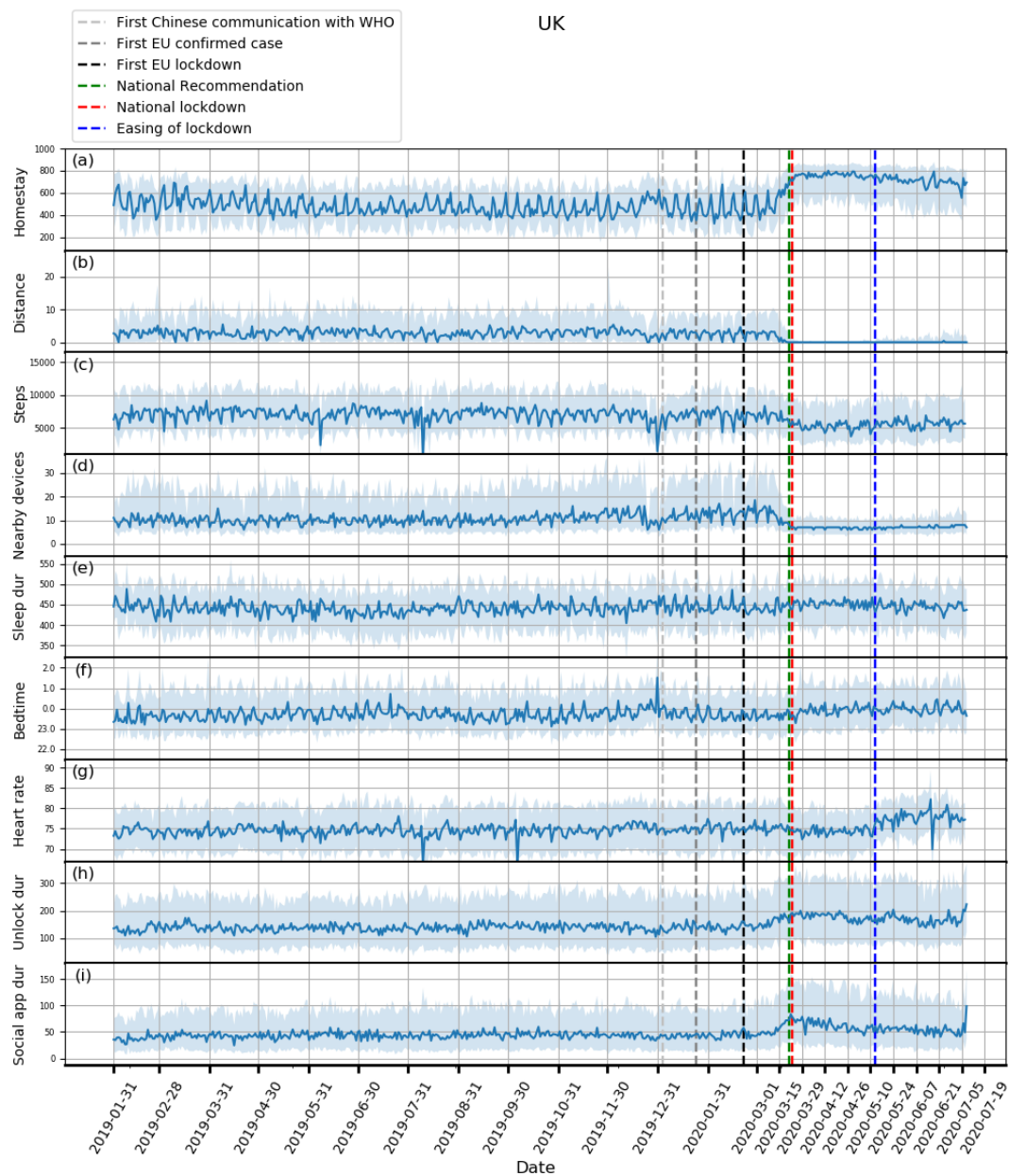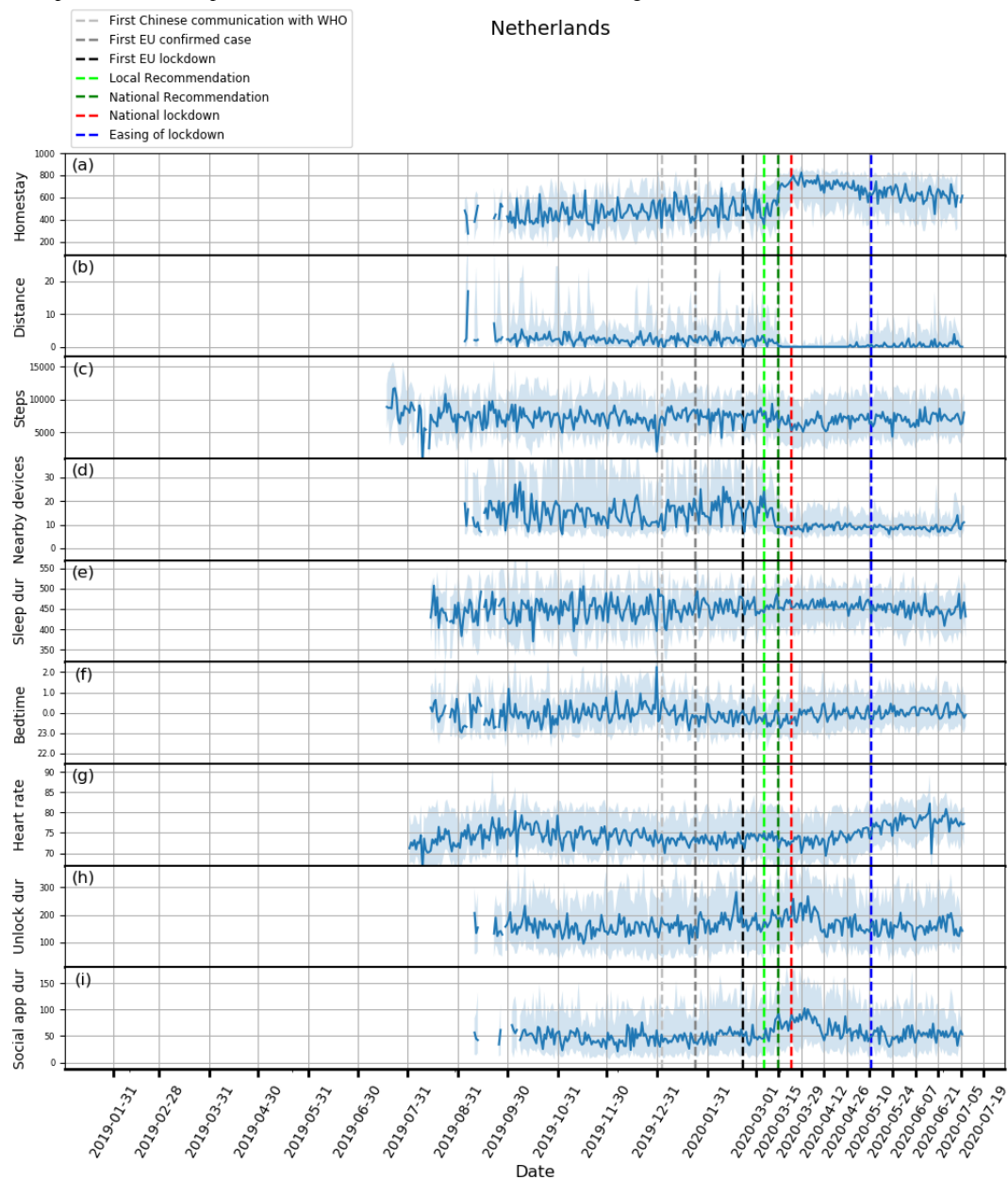
Figure 4. Behavioral changes for the United Kingdom (316 participants). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.

Figure 5. Behavioral changes for the Netherlands (103 participants). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.

Figure 6. Box plots for comparisons among baseline, prelockdown, and during lockdown phases for different features. (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, and (i) social app duration. *means *P*<.05, **means *P*<.01, ***means *P*<.001.
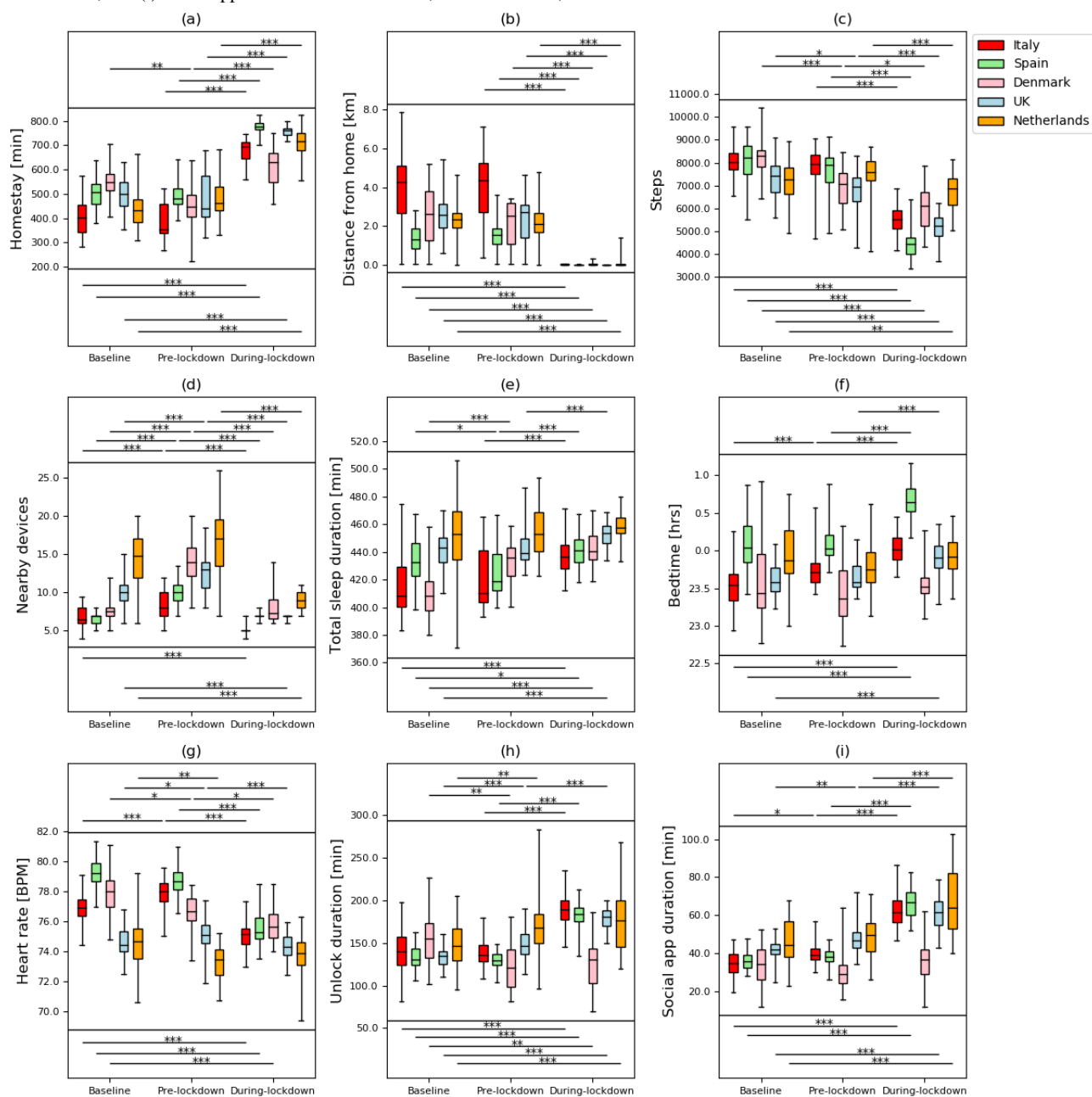
Table 2. Results of post hoc Dunn test (after Kruskal-Wallis tests) on the extracted features between prelockdown and during lockdown periods (only statistically significant differences were reported).

Features	Italy	Spain	Denmark	UK	Netherlands
Homestay					
z test	–9.38	–8.98	–5.44	–9.19	–7.33
P value	<.001	<.001	<.001	<.001	<.001
Maximum distance from home					
z test	9.0	8.91	5.48	8.40	7.58
P value	<.001	<.001	<.001	<.001	<.001
Steps					
z test	8.23	7.72	2.57	6.82	4.78
P value	<.001	<.001	.02	<.001	<.001
Maximum number of nearby devices					
z test	9.68	8.16	5.06	10.2	7.73
P value	<.001	<.001	<.001	<.001	<.001
Total sleep duration					
z test	–4.65	–5.23	__[a]	–4.24	—
P value	<.001	<.001	—	<.001	—
Bedtime					
z test	–4.31	–7.23	—	–5.28	—
P value	<.001	<.001	—	<.001	—
Heart rate					
z test	9.94	7.61	2.68	4.18	—
P value	<.001	<.001	.02	<.001	—
Unlock duration					
z test	–8.8	–8.57	—	–6.0	—
P value	<.001	<.001	—	<.001	—
Social app use duration					
z test	–7.72	–2.36	—	–5.72	–4.98
P value	<.001	<.001	—	<.001	<.001

[a]The difference was not statistically significant.

Figure 7. Zoomed in time series plots for Copenhagen, Denmark (left) and the United Kingdom (right). (a) homestay duration, (b) maximum distance from home, (c) Fitbit step count, (d) maximum number of nearby devices, (e) total sleep duration, (f) bedtime, (g) heart rate, (h) unlock duration, (i) social app duration, (j) COVID-19 confirmed and death cases. Solid line: median; shade: 25th percentile to 75th percentile. dur: duration; WHO: World Health Organization.
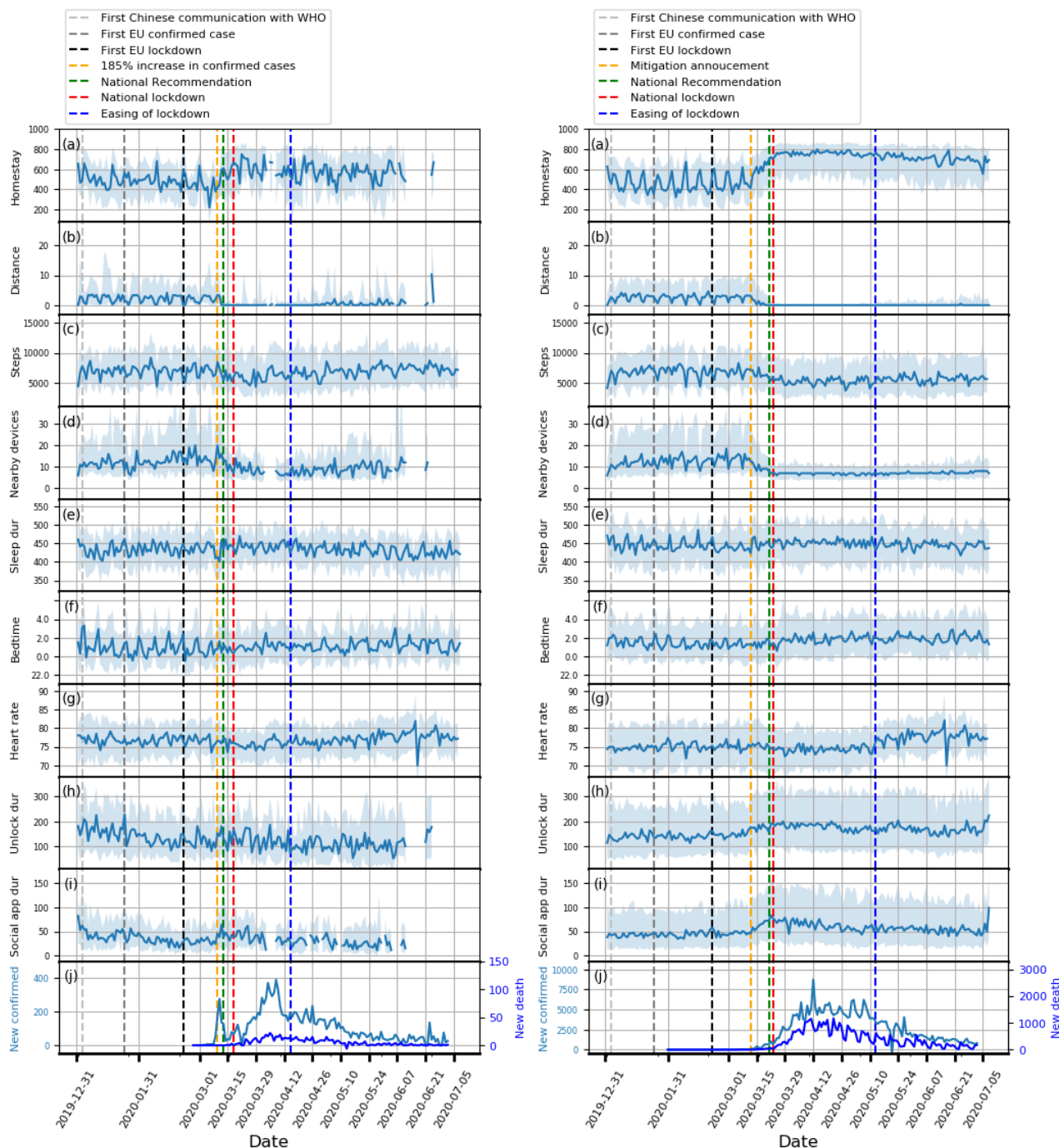
The differences across countries existed in the implemented NPIs as well. The requirement of staying at home except for essential trips and the cancellation of public events were implemented in all countries but Denmark where they were only recommended. Working places were required to close for some sectors in Spain, the United Kingdom, and Denmark, and were required to close for all but essential works in the Netherlands and Italy. Public transport was recommended to close in Italy, Spain, and Denmark. Among all countries, Spain had the least strict restrictions on gatherings and school closures (only geographically targeted).

We observed that the young group spent more time at home in Italy, Spain, and the United Kingdom, and degree holders spent more time at home in Italy and Denmark. The young group took fewer daily steps in Italy, the United Kingdom, and the Netherlands; the low BMI group took fewer daily steps in Italy, Spain, Denmark, and the United Kingdom; the young men group took fewer daily steps in Italy, the United Kingdom, and the Netherlands. Participants educated to degree level walked more in the United Kingdom and the Netherlands but less in Italy. The detailed results are presented in Table 3.

Table 3. Wilcoxon signed rank test results on the examined factors during lockdown periods (only statistically significant differences were reported).

Features and factors	Italy	Spain	Denmark	UK	Netherlands
Homestay					
Age					
w test	108	252	—[a]	3	—
P value	<.001	.003	—	<.001	—
Gender					
w test	—	—	—	—	—
P value	—	—	—	—	—
Degree					
w test	270	—	0	—	—
P value	<.001	—	.004	—	—
BMI					
w test	—	—	—	220	—
P value	—	—	—	<.001	—
Young men					
w test	—	—	—	—	—
P value	—	—	—	—	—
Step					
Age					
w test	283.5	—	—	285.5	151
P value	<.001	—	—	.004	<.001
Gender					
w test	67.5	277	—	—	—
P value	<.001	.005	—	—	—
Degree					
w test	76	—	—	6	0
P value	<.001	—	—	<.001	<.001
BMI					
w test	0	70	38	17	—
P value	<.001	<.001	.001	<.001	—
Young men					
w test	149	—	26	—	0
P value	<.001	—	<.001	—	.02

[a]The difference was not statistically significant.

Discussion

Principal Findings

We quantitatively investigated COVID-19 and associated lockdown-related changes in mobility, physiological measures, and phone use features derived from passive data collected through mobile devices (smartphones and wearable Fitbit devices) of participants recruited in five European countries to the RADAR-CNS program. We were able to measure significant changes in behavioral features between baseline, prelockdown, and during lockdown periods. As well as confirming expected changes such as spending more time at home, travelling much less, having far fewer nearby devices, we observed that people were more active on their phones, interacting with others through social apps particularly around major news events such as national lockdown, suggesting physical but maybe less social distancing. Furthermore, participants had lower heart rates, slept more, and went to bed later. In addition, we found that younger people spent more time at home and took fewer daily steps. Participants with lower BMI took more steps while maintaining comparable homestay with the higher BMI group. With 5 billion

global smartphone users and 500 million smartwatch and wearable device users [28,29], we propose that the ability to generate metrics such as these is vital for evaluating NPIs efficacy.

Our mobility analyses are in line with Google mobility reports [30], where substantial reductions in mobility and increase in residential stays during lockdown periods were found in Italy, Spain, and the United Kingdom; Denmark and the Netherlands by comparison showed an increase in mobility trends for parks and a relatively small increase in residential stays. However, in comparison to Google mobility reports, which provide valuable aggregated data for short periods, RADAR-base is an open-source highly configurable platform that supports collection and analysis of participant-level mobile and phone data in near real time with a potential for targeted interventions. Specifically, focused test and tracing may be directed to people perceived to be at high risk based on their behavior. In addition, RADAR-base was also used to collect self-reported questionnaires related to emotional well-being, functional status, and disease symptom severity of its participants [19]. Since April 2020, new questionnaires have been distributed to specifically assess COVID-19 symptoms and diagnosis status of our RADAR-CNS research participants. Our future work will use the entirety of these data to investigate the potential of wearable data, such as digital early warning signs of COVID-19, the impact of COVID-19 on the QoL, and the clinical trajectories of their primary diagnosis (MDD or MS).

The difference in the response across nations may reflect differences in the implementation of NPIs, media communication, and cultural differences. Denmark implemented stricter restrictions on working places and public transports but were less strict on homestay and public events [24]. In contrast, Spain was more flexible on restrictions of group gatherings and school closures. The contrast in the implementation of different NPIs between the two countries showing distinct behaviors during lockdown sheds light on which NPIs might be more productive in promoting physical distancing. This shows the potential utility of RADAR-base for remotely monitoring the effect of different NPIs, and we also saw evidence of this in our data with participants beginning to return to prelockdown routines as NPIs were lifted. Future work will compare these differences within a country and across countries, which may further elaborate on the effect and impact of NPIs on infection rates and potential second waves.

It is interesting to note that the younger group in general stayed at home more and took fewer steps than the older group. Since most countries required staying at home except for essential trips, one reason could be that the older adults, often less experienced in using online shopping, had to go out for groceries. This conjecture requires future work to investigate. Those educated to the degree level stayed at home for significantly longer in Italy and Denmark, possibly reflecting higher employment in sectors better able to work from home. The low BMI group took more steps but retained similar homestay to the high BMI group, which suggested they may have found other means to exercise locally. This information helps us understand the effectiveness of the NPIs at a

subpopulation level and may be useful in informed strategies for targeted NPIs.

The ability to simultaneously manage multiple data modalities in RADAR-base facilitates the joint analysis and interpretation of them together with NPIs. The decrease in heart rate may be explained by the concurrent reduction in steps, the increase in homestay, and total sleep duration. The reduction in mobility, coupled with an increase in phone use, could possibly serve as indicators of physical distancing observance and resultant compensated social interaction. The delayed bedtime might be related to children homeschooling as a result of school closure, increased phone use, and a lack of exercise. As such, RADAR-base can also be applied to monitor the population health when jointly interpreting features such as step count, sleep duration, and bedtime, which is vital if the social distancing is implemented for a longer duration.

Finally, it has been shown that an elevated resting heart rate may suggest acute respiratory infections [31]. It may be possible to infer one's infection by continuously monitoring heart rate, especially when the population remains indoor for a vast majority of the time. Such monitoring provides the possibility to generate early warning signals for symptomatic or presymptomatic respiratory infections, thereby aiding timely self-isolation or treatment. The COVID-19–related symptom and diagnosis questionnaires have been added to the study and may provide a means to investigate these relationships further.

Media Effects

In addition to changes in trends over longer periods, we also identified interesting findings in relation to specific events (see Figure 7). A dramatic reduction in total sleep duration was observed in Denmark around March 11, 2020, which may be related to the announcement of the pending lockdown on that day and a 185% increase in the confirmed cases in Denmark on the previous day. Another example can be seen just after the mitigation phase was announced in the United Kingdom on March 12, in which social distancing was not strongly recommended, yet we saw participants isolating themselves voluntarily by staying at home for much longer. These observations highlight the potential role of media and social media in the distribution of information that may precipitate certain behavior. This observation may also explain the significant difference between the baseline and prelockdown phases, and suggests that people may have acted ahead of further government restriction. Furthermore, this is accompanied by a marked loss of week day and weekend periodic structure prelockdown and during the lockdown periods.

Limitations

There are some issues to consider concerning this work. First, the participants included in this study have different medical conditions (MDD or MS), which led to different baseline levels across countries. Nevertheless, as the focus of this study is the changes in the prelockdown and during lockdown phases relative to the baseline, we were still able to identify and compare the changes induced by lockdowns. We also analyzed the data collected in Spain split into MDD and MS separately and found the results only differed marginally. Understanding of any

artifacts or effects introduced into the RADAR-CNS data by the NPIs will be crucial in RADAR-CNS being able to deliver its aim of identifying signals that predict and prevent MDD and MS. Although the medical conditions of the population in this study might not be fully generalizable to a wider population healthy or with other conditions, we were able to demonstrate the utility of RADAR-base in monitoring behavioral changes, which can be readily generalized to other cohorts.

Second, the individual disease status at baseline may be different from that of the during lockdown period in each country, which might complicate the comparisons. To mitigate this, we used the same time period from the previous year to suppress the seasonal variability. We believe this complication on the population level was unlikely to be large, especially compared to the impact of lockdowns.

Third, participants recruited at different times may use different devices for smartphones and Fitbit depending on the availability and enrollment dates, which might make it difficult for interparticipant comparisons. Yet, this work focused on the population-level behavioral changes induced by NPIs where the handset variability was less of a factor.

Fourth, on account of requirements for participants' privacy in the RADAR-CNS studies, location data were purposely obfuscated with a participant-specific random value preventing precise localization of the participants, which limited use of regional geographic factors within a country. It would be interesting to examine how specific regions react to lockdowns when these data are available in future work.

Fifth, limited sample sizes in certain countries and data loss impacted the smoothness of the plots showing how the extracted features evolved over time. The plots for Denmark and the Netherlands showed relatively large variance particularly in the early phase, as these sites have only recently begun recruiting. Several dips and spikes in step counts and heart rate were seen in all countries during July and August. This observation was due to having data loss because of connectivity issues with the Fitbit server during this time.

Last, we only explored a subset of features that can be derived from smartphones and Fitbit wearable devices. Future work will investigate whether other features offer additional information for a more complete description of lifestyle changes.

Conclusions

Using participants' data from smartphones and wearable devices collected and managed by RADAR-base over 1.5 years covering the outbreak and subsequent spread of the COVID-19 pandemic across five European countries, we were able to detect and monitor the physical-behavioral and social-behavioral changes in response to the NPIs. We found that, as well as expected findings (that validated the data collection platform) relating to increased time spent at home, less travel, and fewer nearby Bluetooth-enabled devices, participants were more active on their phone, in particular, interacting with others using social apps, especially around major news events, suggesting increased physical distancing with socialising and interaction moving online. Furthermore, we found that participants had lower heart rates, slept more, and went to bed later. We demonstrated different responses across countries with Denmark showing attenuated responses to NPIs compared to other countries, which may be associated with their different focus of implementation NPIs. We found that younger people stayed at home for longer yet walked less compared to older adults and that the people with lower BMI remained more active during lockdown while having comparable homestay compared to their counterparts with higher BMI. Joint analysis of the extracted features is important for evaluating aspects of NPIs performance during their introduction and any subsequent relaxation of these measures. This work demonstrates the value of RADAR-base for collecting data from wearables and mobile technologies to understand the effect and response of public health interventions implemented in response to infectious outbreaks such as COVID-19. This ability to monitor response to interventions, in near real time, will be particularly important in understanding behavior as social distancing measures are relaxed as part of any COVID-19 exit strategy. Future work will include using participants' responses to COVID-19–related questionnaires together with an expanded feature set to gain more specific understandings into the relationship between mobile devices–derived features and the COVID-19 symptoms.

XSL•FO

RenderX

Authors' Contributions

SS (Shaoxiong Sun), AAF, and RJBD contributed to the study design. SS (Shaoxiong Sun) contributed to the data analysis, figures drawing, and manuscript writing. AAF, NC, TW, VAN, GC, MH, and RJBD contributed to the critical revision of the manuscript. AAF, YR, ZR, PC, CS, and RJBD contributed to the platform design and implementation. AAF, IMG, AR, TW, VAN, GC, MH, and RJBD contributed to the administrative, technical, and clinical support of the study. FM, GDC, SS (Sara Simblett), LL, ALG, AZ, BWJHP, FL, SS (Sara Siddi), and JMH contributed to data collectionn.

Conflicts of Interest

VAN is an employee of Janssen Research & Development LLC and may own equity in the company.

References

1. Bai Y, Yao L, Wei T, Tian F, Jin D, Chen L, et al. Presumed asymptomatic carrier transmission of COVID-19. JAMA 2020 Feb 21:1406-1407 [FREE Full text] [doi: 10.1001/jama.2020.2565] [Medline: 32083643]
2. Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. Lancet Respir Med 2020 Apr;8(4):420-422. [doi: 10.1016/s2213-2600(20)30076-x]
3. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J Travel Med 2020 Mar 13;27(2):taaa021 [FREE Full text] [doi: 10.1093/jtm/taaa021] [Medline: 32052846]
4. WHO coronavirus disease (COVID-19) dashboard. World Health Organization. 2020. URL: https://covid19.who.int/ [accessed 2020-07-01]
5. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020 Feb;395(10223):497-506. [doi: 10.1016/s0140-6736(20)30183-5]
6. Yu P, Zhu J, Zhang Z, Han Y. A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period. J Infect Dis 2020 May 11;221(11):1757-1761 [FREE Full text] [doi: 10.1093/infdis/jiaa077] [Medline: 32067043]
7. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
8. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020 Feb 24:A. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]
9. Lau H, Khosrawipour V, Kocbach P, Mikolajczyk A, Schubert J, Bania J, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J Travel Med 2020 May 18;27(3):taaa037 [FREE Full text] [doi: 10.1093/jtm/taaa037] [Medline: 32181488]
10. Lazzerini M, Putoto G. COVID-19 in Italy: momentous decisions and many uncertainties. Lancet Global Health 2020 May;8(5):e641-e642. [doi: 10.1016/s2214-109x(20)30110-8]
11. World economic outlook update. International Monetary Fund. 2020 Jun. URL: https://www.imf.org/~/media/Files/Publications/WEO/2020/Update/June/English/WEOENG202006.ashx?la=en [accessed 2020-07-18]
12. Pandian P, Mohanavelu K, Safeer K, Kotresh T, Shakunthala D, Gopal P, et al. Smart Vest: wearable multi-parameter remote physiological monitoring system. Med Eng Phys 2008 May;30(4):466-477. [doi: 10.1016/j.medengphy.2007.05.014] [Medline: 17869159]
13. Cancela J, Pastorino M, Tzallas A, Tsipouras M, Rigas G, Arredondo M, et al. Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. Sensors (Basel) 2014 Sep 16;14(9):17235-17255 [FREE Full text] [doi: 10.3390/s140917235] [Medline: 25230307]
14. Paradiso R, Alonso A, Cianflone D, Milsis A, Vavouras T, Malliopoulos C. Remote health monitoring with wearable non-invasive mobile system: The Healthwear project. In: IEEE. 2008 Oct 14 Presented at: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 20-25 Aug 2008; Vancouver, BC. [doi: 10.1109/iembs.2008.4649503]
15. RADAR-base. URL: https://radar-base.org/
16. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Hyve, RADAR-CNS Consortium. RADAR-Base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR mHealth uHealth 2019 Aug 01;7(8):e11734 [FREE Full text] [doi: 10.2196/11734] [Medline: 31373275]
17. Stewart C, Rashid Z, Ranjan Y, Sun S, Dobson R, Folarin A. RADAR-base: major depressive disorder and epilepsy case studies. In: UbiComp '18: Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. 2018 Oct Presented at: UbiComp '18: The 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing; October, 2018; Singapore, Singapore p. 1735-1743. [doi: 10.1145/3267305.3267540]
18. RADAR-CNS. URL: https://radar-cns.org/

XSL•FO
RenderX

19.   Matcham F, Barattieri di San Pietro C, Bulgari V, de Girolamo G, Dobson R, Eriksson H, RADAR-CNS consortium. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatry 2019 Feb 18;19(1):72 [FREE Full text] [doi: 10.1186/s12888-019-2049-z] [Medline: 30777041]

20.   hersenonderzoek.nl. URL: https://hersenonderzoek.nl/

21.   Ester M, Kriegel H, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD-96 Proceedings. 1996 Presented at: Second International Conference on Knowledge Discovery and Data Mining; 1996 Aug 2-4; Portland, OR p. 226-231.

22.   Sleep logs. Fitbit SDK. URL: https://dev.fitbit.com/build/reference/web-api/sleep/ [accessed 2020-07-18]

23.   Communication. Google Play. URL: https://play.google.com/store/apps/category/COMMUNICATION?hl=en_GB [accessed 2020-07-18]

24.   Hale T, Petherick A, Phillips T, Webster S, Kira K, Angrist N. Coronavirus government response tracker. Blavatnik School of Government. 2020 Mar. URL: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker [accessed 2020-04-16]

25.   Ostertagová E, Ostertag O, Kováč J. Methodology and application of the Kruskal-Wallis test. AMM 2014 Aug;611:115-120. [doi: 10.4028/www.scientific.net/amm.611.115]

26.   Dinno A. Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. Stata J 2018 Nov 19;15(1):292-300. [doi: 10.1177/1536867x1501500117]

27.   Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Statistics 2001 Aug;29(4):1165-1188. [doi: 10.1214/aos/1013699998]

28.   Number of mobile phone users worldwide from 2015 to 2020. Statista. URL: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/ [accessed 2020-07-18]

29.   Number of connected wearable devices worldwide by region from 2015 to 2022. Statista. URL: https://www.statista.com/statistics/490231/wearable-devices-worldwide-by-region/ [accessed 2020-07-18]

30.   COVID-19 community mobile reports. Google. URL: https://www.google.com/covid19/mobility/ [accessed 2020-07-18]

31.   Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. Lancet Digital Health 2020 Feb;2(2):e85-e93. [doi: 10.1016/s2589-7500(19)30222-5]

Abbreviations

EFPIA: European Federation of Pharmaceutical Industries and Associations
IMI2: Innovative Medicines Initiative
MDD: major depressive disorder
MS: multiple sclerosis
NIHR: National Institute for Health Research
NPI: nonpharmaceutical intervention
QoL: quality of life
RADAR: Remote Assessment of Disease and Relapse
RADAR-CNS: Remote Assessment of Disease and Relapse–Central Nervous System