Differential effects of COVID-19 lockdowns on well-being: interaction between age, gender and chronotype

Shay Oved¹⁺, Merav Mofaz¹⁺, Anat Lan²⁺, Haim Einat², Noga Kronfeld-Schor³, Dan Yamin¹⁴, and Erez Shmueli¹*¹

¹Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel
²School of Behavioral Sciences, The Academic College of Tel Aviv-Yafo, Tel-Aviv, Israel
³School of Zoology and Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv, Israel
⁴Center for Combating Pandemics, Tel-Aviv University, Tel-Aviv, Israel
⁺Equal contribution
*Corresponding author: shmueli@tau.ac.il

ABSTRACT

The unprecedented restrictions imposed due to the COVID-19 pandemic, including movement control orders and lockdowns, altered our daily habits, and severely affected our well-being and physiology. The effect of these changes is yet to be fully understood. Here, we analyzed highly detailed data on 169 participants for 2-6 months, before and during the second COVID-19 lockdown in Israel. Our entire study was conducted during the COVID-19 pandemic, and therefore is the first to decipher the specific effects of the lockdown from the general effects of the pandemic. We extracted 12 well-being indicators from sensory data of smartwatches and from self-reported questionnaires, filled on a daily basis using a designated mobile application. We used a mixed ANOVA model to study the interplay between age, gender, and chorotype on well-being before and after lockdowns. We found that at the population level, lockdowns resulted in significant changes in mood, sleep duration, sport duration, social encounters, resting heart rate, and the number of steps. The lockdown's adverse effects were greater for young early chronotypes who did not increase their sleep duration, reduced activity level and suffered from significantly reduced mood, and for women, who further suffered an increase in stress levels and a greater decline in social encounters. Our findings underscore that while lockdowns severely impacted our well-being and physiology in general, greater damage has been identified in certain subpopulations. Based on the observed effects, special attention should be given to younger people, who are usually not in the focus of social support, and to women.

1 Introduction

Restrictions against the COVID-19 pandemic have altered everyday life in many ways. Our daily habits – the time we go to sleep and wake up, eat, commute, work, and engage in social activities – all changed as part of the battle against the disease, with many countries imposing social distancing regulations and mobility restrictions such as lockdowns¹³. These changes are expected to affect our well-being and physiology. Importantly, in the context of contracting and fighting viral infections, these changes can also weaken our immune system. Thus, it is imperative to understand the complex consequences of social distancing and lockdowns.

Recent studies exploring the influence of COVID-19 on well-being found both negative and positive effects. Some studies found that lockdowns were associated with adverse outcomes such as high levels of anxiety and stress, and lower physical activity⁴–⁶. In contrast, lockdowns were also correlated with favorable changes in sleep patterns, including longer sleep duration⁷–¹⁰. Moreover, some of the documented effects of lockdown are inconsistent. For example, Leone found a positive, albeit small, impact of lockdowns on sleep quality¹⁰, whereas others reported an adverse opposite change⁵,⁸,¹¹, or no change at all¹². Such discrepancies could be explained by variability in the response of subpopulations.

Better understanding of the different effects of social distancing and lockdowns and their consequences on diverse populations can have a critical impact on the recommended behavior at the individual level, with the possibility of specific directions, instructions, medical advice and social care to separate sub-groups. Furthermore, it may also bare important insights for our “new normal”, post-COVID-19 life. Only a few studies thus far examined the differential effects of the social changes induced by the disease between subpopulations. Several studies suggested more profound effects in younger adults. Specifically, lockdowns were associated with higher depressive symptoms, stress and anxiety, lower sleep quality levels, longer sleep duration, and later sleep onset and offset in younger compared to older participants¹⁰,¹²–¹⁴. Several other studies reported
A growing body of literature has found associations between a later chronotype and increased risk for depression and mood disorders, higher metabolic dysfunction rates and morbidity, and reduced physical activity. COVID-19 lockdowns have changed both our social schedules and the synchronizers of our biological clock, including light exposure: most of us spend less time outdoors exposed to sunlight during the day, but are exposed to artificial light during the night. Such changes are expected to affect different chronotypes differently. Indeed, a recent study using self-reported questionnaires found an influence of age and chronotype on the magnitude of the delay in sleep onset time in response to lockdowns.

The COVID-19 pandemic provides a unique opportunity to test the effects of social and mobility changes on well-being. Importantly, these effects may also indirectly affect the immune system: sleep, biological rhythms and well-being are all interconnected, and affect our immune system function and ability to fight viral infections. Therefore, understanding these relations may also help us develop protective recommendations for reducing the risk of infection and improve vaccination outcomes, which may be different for different subpopulations.

We specifically sought to probe whether age, gender and chronotype moderate the effects of such changes on well-being using both wearable monitoring devices (Fitbit) and daily self-reported digital questionnaires. Data were collected from 169 participants between May 11 to October 17 (2020), before and during the second COVID-19 lockdown in Israel. We used these data to evaluate the change in various well-being indicators including sleep patterns, physical activity and mood.

2 Methods

2.1 The PerMed Pilot Study

The source of the primary data for this study comes from the PerMed Study. The PerMed study was approved by Tel-Aviv University Institutional Review Board (IRB) and is conducted under strict protocol guidelines. The goal of the PerMed study is to improve diagnosis of highly common acute infectious diseases through the combined use of Electronic Medical Records (EMR) and behavioral information collected from smartphones and smartwatches. The study aims to perform an observational study to collect data from a cohort of 5,000 participants for two years.

For that purpose, we developed a dedicated data collection platform. One of the main components of this platform is the PerMed mobile application that passively collects smartphone sensors data, as well as self-reported daily questionnaires.

As a preparation to the PerMed study, we performed a smaller and shorter pilot study with 192 participants for 2-6 months. The objective of this pilot was to evaluate the performance of the data collection platform and the process of recruiting and engaging participants. Participants in the pilot study were equipped with our dedicated mobile application, and a Fitbit Inspire HR smartwatch. In the current study, we analyze data that were collected as part of the PerMed pilot study to test the effects of the second COVID-19 lockdown in Israel on various indicators of well-being in different subpopulation.

In the following subsections we provide more details on how the participants were recruited, the time periods of the study, and the collected data.

2.1.1 Subjects recruitment

Subjects were recruited through advertisements on social media. After signing a consent form, they filled an enrollment questionnaire, had the mobile study apps installed on their smartphone and received the smartwatch. Overall, we recruited 192 Israeli participants above the age of 18. Additional details about the recruitment and engagement process of participants, are provided in Appendix A.
2.1.2 Time periods
The PerMed Pilot Study started on May 11, 2020. In this study, we examined data that were collected until October 19, 2020.

2.1.3 Data collection
In this study, we analyzed two sources of data that were collected as part of the PerMed pilot study:

- The daily questionnaire consisted of eight questions about mood, stress, sport duration, sleep duration, sleep quality, encounters with other people, clinical symptoms and diagnosed diseases. A detailed description of the questionnaire is provided in Appendix B.

- Smartwatch data were collected at a daily aggregation level, and consisted of average resting heart rate, active minutes, steps, distance, calories and sleep data, including bed times, wake times, time in bed and total sleep duration.

2.2 Experimental Setting
We evaluated the effects of the second COVID-19 lockdown in Israel on different subpopulations for various well-being indicators. For this purpose, we analyzed data collected as part of the PerMed pilot study, which included daily questionnaires and smartwatch data. In the following subsections, we provide more details on the considered time periods, how data were preprocessed, and how they were analyzed.

2.2.1 Time periods
We considered two time periods: (1) the period from the beginning of the PerMed pilot study - May 11, 2020, and a week before the beginning of the second COVID-19 lockdown in Israel - September 11, 2020, and (2) the lockdown period - September 18, 2020 to October 17, 2020. During the week between September 12, 2020 and September 18, 2020, the intention to conduct a national lockdown has been widely discussed in the media. Thus, to neutralize the announcement and discussion’s side effects, we did not analyze data from this specific week.

2.2.2 Data preprocessing
Before analyzing the data, we performed several preprocessing steps. First, in cases where participants filled the daily questionnaire more than once on the same day, only the latest answers for that day were considered.

Next, sleep records were processed to identify the main sleep intervals. More specifically, in the vast majority of cases participants had at most one sleep record per day. However, in roughly 20% of the cases, sleep data collected by the smartwatch included more than one sleep record for the same participant for the same day. This could be due to participants waking up once or more during their main sleep, or if they had afternoon naps in addition to their main sleep. To identify the segments belonging to the main sleep, we first merged consecutive sleep records with a difference of up to two hours between them into a single sleep interval. Then, sleep intervals with a total duration of fewer than three hours or start time between 10 AM and 4 PM were omitted.

Then, for each participant, we calculated nine well-being indicators. The first six indicators come from the daily self-reported questionnaire. They include: mood (on a scale of -2 to 2, where -2 means awful and 2 means excellent), stress (on a scale of -2 to 2, where -2 means very low and 2 means very high), sleep quality (on a scale of -2 to 2, where -2 means awful and 2 means excellent), sleep duration (hours), sport duration (minutes), and the number of social encounters. The other three indicators were extracted from the smartwatch, and include: resting heart rate (bpm), steps, and sleep duration (hours). We calculated the mean value for each participant and each of the nine indicators, before and during the lockdown.

In order to capture changes in sleep routines, we also calculated the following three indicators:

1. Mid sleep point free days (MSF) - For each participant and for each week, we calculated the weekly MSF value as the average of middle sleep time in free days (weekends and public holidays). Then, the MSF value for that participant was calculated as the average over all weekly MSF values - once for the period before the lockdown and once for the period during the lockdown. For the calculation, Jewish weekends (Friday and Saturday) and holidays were marked as free days.

2. Mid sleep point work days (MSW) - For each participant and for each week, we calculated the weekly MSW value as the average of middle sleep time in work days. Then, the MSW value for that participant was calculated as the average over all weekly MSW values - once for the period before the lockdown and once for the period during the lockdown. The calculation of MSW for a given week was done only if the participant had at least two sleep records in this week’s workdays.

3. Social Jetlag (SJL) - For each participant and for each week that had a weekly MSF value and a weekly MSW value, we calculated the weekly SJL value as the absolute difference between them. Then, the SJL value for that participant was calculated as the average over all weekly SJL values - once for the period before the lockdown and once for the period during the lockdown.
Finally, 23 of 192 participants were omitted from the analysis because they lacked MSF information in the period before lockdown (most probably because they barely wore their smartwatch at nights or weekends).

### 2.2.3 Statistical Analyses

In order to test the effect of the lockdown on different subpopulations for different well-being indicators, we applied a Mixed ANOVA test. More specifically, we considered the 12 well-being indicators mentioned in the previous subsection. For each of these 12 indicators, we used a separate Mixed ANOVA test, where the dependent variable was the indicator, and the four independent variables (main factors) were:

1. **Lockdown** - a within-subjects factor with two levels: Before Lockdown and During Lockdown.

2. **Age Group** - a between-subjects factor with two levels: Younger and Older. The groups were divided based on the median value.

3. **Gender** - a between-subjects factor with two levels: Men and Women.

4. **Chronotype** - a between-subjects factor with two levels: Early chronotype and Late chronotype. The groups were divided based on the median value of MSF before the lockdown.

More formally, for each of the 12 well-being indicators, the considered Mixed ANOVA model includes the four main factors and all of their potential interactions:

$$Indicator \sim Lockdown \ast Age\_Group \ast Gender \ast Chronotype$$  \hspace{1cm} (1)

We performed post-hoc multiple comparison Bonferroni tests for the significant interactions, and measured Cohen’s d effect size for the significant effects.

### 3 Results

#### 3.1 Descriptive statistics

The age of the participants ranged between 20 and 80, with clear two main age groups: 20-40 and 60-80. Out of the 169 participants, 94 (55.62%) were women, and 75 (44.38%) were men. As captured by their MSF in the period before the lockdown, the chronotypes of participants were characterized by a Gaussian shape, with a mean of 4:05 AM local time (Fig. 1).

As can be seen from Fig. 1: (1) the two age groups distribute roughly evenly between men and women (Fig. 1A); (2) the younger age groups is mainly associated with late chronotypes while the older age group is mainly associated with early chronotypes (Fig. 1B); and (3) women and men present a similar distribution of chronotypes (Fig. 1C).
Table 1 presents information about the mean and 95% confidence interval for each of the 12 examined well-being indicators, for the time periods before and during the lockdown, for the entire population of 169 participants. As can be seen from the table, some indicators (sleep duration, MSF, and MSW) seem to increase during the lockdown, while some other indicators (mood, encounters, resting heart rate, steps) seem to decrease. A breakdown of this table according to age group, gender and chronotype is provided in Appendix C.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Before Lockdown</th>
<th>During Lockdown</th>
<th>Difference</th>
<th>CI LB</th>
<th>CI UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood</td>
<td>0.86</td>
<td>0.78</td>
<td>-0.08</td>
<td>-0.13</td>
<td>-0.02</td>
</tr>
<tr>
<td>Stress</td>
<td>-0.69</td>
<td>-0.69</td>
<td>0.00</td>
<td>-0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Sleep quality</td>
<td>0.47</td>
<td>0.49</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>Sleep duration (hours)</td>
<td>6.29</td>
<td>6.42</td>
<td>0.13</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>Sport duration (minutes)</td>
<td>31.34</td>
<td>29.04</td>
<td>-2.31</td>
<td>-4.64</td>
<td>0.03</td>
</tr>
<tr>
<td>Encounters</td>
<td>12.09</td>
<td>7.86</td>
<td>-4.23</td>
<td>-5.36</td>
<td>-3.10</td>
</tr>
<tr>
<td>Resting heart rate (bpm)</td>
<td>62.90</td>
<td>62.23</td>
<td>-0.67</td>
<td>-1.11</td>
<td>-0.23</td>
</tr>
<tr>
<td>Steps</td>
<td>8671.11</td>
<td>7967.67</td>
<td>-703.44</td>
<td>-1032.90</td>
<td>-373.98</td>
</tr>
<tr>
<td>Sleep duration (hours)</td>
<td>7.01</td>
<td>7.15</td>
<td>0.14</td>
<td>0.05</td>
<td>0.22</td>
</tr>
<tr>
<td>MSF (local time)</td>
<td>4.05</td>
<td>4.16</td>
<td>0.19</td>
<td>0.07</td>
<td>0.31</td>
</tr>
<tr>
<td>MSW (local time)</td>
<td>3.41</td>
<td>3.48</td>
<td>0.12</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>SJL (hours)</td>
<td>0.84</td>
<td>0.86</td>
<td>0.02</td>
<td>-0.09</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Table 1. Descriptive statistics. Each row represents a single well-being indicator and contains: the mean value before lockdown, the mean value during lockdown, the mean difference, and the 95% confidence interval (lower bound and upper bound) of the difference.
3.2 Within-subjects effects

Table 2 presents the p-values of the within-subjects effects for the Mixed ANOVA test that was conducted for each of the 12 examined well-being indicators. Each row represents a single indicator (and thus also a single test) and each column represents the examined factor/interaction. Each entry represents the p-value of the corresponding effect, where statistically significant effects are marked with asterisks. A complementary table of the between subject effects is provided in Appendix D.

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Indicator</th>
<th>Lockdown</th>
<th>Lockdown* Age Group</th>
<th>Lockdown* Gender</th>
<th>Lockdown* Chronotype</th>
<th>Lockdown* Age Group* Gender</th>
<th>Lockdown* Gender* Chronotype</th>
<th>Lockdown* Age Group* Gender* Chronotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood</td>
<td></td>
<td>0.00***</td>
<td>0.05**</td>
<td>0.58</td>
<td>0.58</td>
<td>0.14</td>
<td>0.02**</td>
<td>0.80</td>
</tr>
<tr>
<td>Stress</td>
<td></td>
<td>0.90</td>
<td>0.36</td>
<td>0.01***</td>
<td>0.30</td>
<td>0.24</td>
<td>0.92</td>
<td>0.73</td>
</tr>
<tr>
<td>Sleep quality</td>
<td></td>
<td>0.69</td>
<td>0.47</td>
<td>0.11</td>
<td>0.16</td>
<td>0.14</td>
<td>0.21</td>
<td>0.16</td>
</tr>
<tr>
<td>Sleep duration</td>
<td></td>
<td>0.00***</td>
<td>0.99</td>
<td>0.42</td>
<td>0.70</td>
<td>0.41</td>
<td>0.71</td>
<td>0.30</td>
</tr>
<tr>
<td>Sport duration</td>
<td></td>
<td>0.08*</td>
<td>0.14</td>
<td>0.45</td>
<td>0.52</td>
<td>0.71</td>
<td>0.89</td>
<td>0.32</td>
</tr>
<tr>
<td>Encounters</td>
<td></td>
<td>0.00***</td>
<td>0.55</td>
<td>0.03**</td>
<td>0.15</td>
<td>0.43</td>
<td>0.10</td>
<td>0.76</td>
</tr>
<tr>
<td>Smartwatch</td>
<td>Resting heart rate</td>
<td>0.05*</td>
<td>0.12</td>
<td>0.06*</td>
<td>0.61</td>
<td>0.23</td>
<td>0.11</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Steps</td>
<td>0.00***</td>
<td>0.01*</td>
<td>0.74</td>
<td>0.41</td>
<td>0.30</td>
<td>0.76</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Sleep duration</td>
<td>0.05**</td>
<td>0.60</td>
<td>0.57</td>
<td>0.35</td>
<td>0.32</td>
<td>0.04**</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>MSF</td>
<td>0.02**</td>
<td>0.11</td>
<td>0.89</td>
<td>0.95</td>
<td>0.23</td>
<td>0.22</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>MSW</td>
<td>0.07*</td>
<td>0.67</td>
<td>0.46</td>
<td>0.28</td>
<td>0.74</td>
<td>0.82</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>S Jill</td>
<td>0.76</td>
<td>0.19</td>
<td>0.80</td>
<td>0.59</td>
<td>0.69</td>
<td>0.16</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 2. Within-subjects effects. Each row represents a Mixed ANOVA test for a single well-being indicator. Each column represents a single factor/interaction, and each entry represents the p-value of the corresponding effect. Statistically significant effects are marked with asterisks.

3.2.1 Effects of the lockdown on the entire population

Overall, we found that the lockdown affected various well-being indicators in the entire population (Table 2). We found a decline in mood (an average level of 0.87 vs. 0.76, p < 0.01, Cohen’s d = 0.14), sport duration (an average number of 30.05 minutes vs. 27.71 minutes, p = 0.08, Cohen’s d = 0.10), encounters (an average number of 11.49 encounters vs. 7.81 encounters, p < 0.01, Cohen’s d = 0.44), resting heart rate (an average of 62.59 bpm vs. 62.11 bpm, p = 0.05, Cohen’s d = 0.09), and steps (an average number of 8453.30 steps vs. 7710.58 steps, p < 0.01, Cohen’s d = 0.19). An increase was observed in sleep duration based on questionnaire responses (an average of 6.28 hours vs. 6.42 hours, p < 0.01, Cohen’s d = 0.14) and based on smartwatch records (an average of 7.03 hours vs. 7.12 hours, p = 0.05, Cohen’s d = 0.14), MSF (an average hour of 4:03am vs. 4:12am, p = 0.02, Cohen’s d = 0.13) and MSW (an average hour of 3:43am vs. 3:49am, p = 0.07, Cohen’s d = 0.08). No significant effects were found for stress, sleep quality and SJL.

3.2.2 Unequal effects of the lockdown on different subpopulations

Interestingly, beyond the general effects, the data clearly demonstrate dissimilar effects in different subpopulations. Moreover, when analyzing separately the different subpopulations, the effect sizes (Cohen’s d) of the differences before and after lockdown become larger. Here, we focus on three different segmentations of the population — age group, gender, chronotype.

A significant interaction was found between Lockdown and Age Group in mood (p = 0.05) and steps (p = 0.01), suggesting a different effect of the lockdown on the two age groups with respect to these two indicators. Specifically, we observe that the negative effect of the lockdown was stronger for the younger age group in both cases (Fig. 2A and 2B). Post-hoc analysis revealed that the decline in mood and steps between the two time periods was significant in the younger age group (p < 0.01, Cohen’s d = 0.24 (mood); p < 0.01, Cohen’s d = 0.35 (steps)) but not in the older age group.
A significant interaction was found between Lockdown and Gender in stress ($p = 0.01$), encounters ($p = 0.03$), and resting heart rate ($p = 0.06$), suggesting a different effect of the lockdown on women and men with respect to these three indicators. Specifically, we observe that men became less stressed during the lockdown while women became more stressed (Fig. 3A); the number of encounters declined for both genders, but for women the decline was sharper (Fig. 3B); and women’s heart rate declined during lockdown while men’s heart rate did not change (Fig. 3C). Post-hoc analysis revealed that the change of stress between the two time periods was significant for women ($p = 0.04$, Cohen’s $d = 0.14$) and men ($p = 0.06$, Cohen’s $d = 0.17$), but the change was in different directions. The change in encounters was also significant for women ($p < 0.01$, Cohen’s $d = 0.56$) and men ($p = 0.02$, Cohen’s $d = 0.30$). The change in resting heart rate was significant for women only ($p < 0.01$, Cohen’s $d = 0.14$).

Figure 2. Effects of the lockdown on different age groups for: (A) mood, and (B) steps. The y-axis represents the Estimated Marginal Means (EMM) for the examined well-being indicator.
We also found a significant interaction between Lockdown, Age Group and Chronotype in mood ($p = 0.02$) and sleep duration recorded by the smartwatch ($p = 0.04$), suggesting a different effect of the lockdown on different combinations of Age Group and Chronotype with respect to these three indicators. Specifically, for the older age group, the decrease in mood was mainly expressed in late chronotypes (Fig. 4A), whereas for the younger age group the decrease in mood was mainly expressed in early chronotypes (Fig. 4B). Similarly, for the older age group, the increase in sleep-duration recorded by the smartwatch was mainly expressed in early chronotypes (Fig. 4C), whereas for the younger age group, it was mainly expressed in late chronotypes (Fig. 4D). Post-hoc analysis revealed that the change of mood between the two time periods was significant in the younger age group for early chronotypes ($p < 0.01$, Cohen's $d = 0.55$). The change in sleep duration was significant or near significant in the younger age group for late chronotypes ($p < 0.01$, Cohen's $d = 0.31$) and in the older age group for early chronotypes ($p = 0.09$, Cohen's $d = 0.12$).
Figure 4. Effects of the lockdown on different combinations of age groups and chronotypes: (A+B) mood, and (C+D) sleep duration recorded by the smartwatch. The y-axis represents the Estimated Marginal Means (EMM) for the examined well-being indicator.
4 Discussion

The current study demonstrates that at the entire population level, the lockdown resulted in lower mood and reduced resting heart rate, fewer social interactions, lower physical activity (number of steps and sport duration), increased sleep duration and later MSF and MSW. We did not find an effect on stress, sleep quality or SJL. Previous studies using surveys comparing control (before the pandemic) and lockdown similarly found that lockdown resulted in longer and later sleep on weekdays, but also reduced social jetlag\textsuperscript{5,10,12}. The increase in sleep duration was significant in our general study population and was evident in both the questionnaire and the smartwatch data. In general, this part of our analysis is clearly in agreement with recent other studies, which were mostly based on self-reported questionnaires.

A more complex and interesting picture appears when we separately examine different subpopulations based on age, gender and chronotype. Specifically, our analyses indicate that sleep duration shows a significant increase in the late chronotypes of the younger group, and a marginally significant increase in the early chronotypes of the older group. Late chronotypes, especially in the age of our younger group which usually work or study and have younger children, normally suffer sleep loss. With no need to wake up the children or go to everyday obligations, the more relaxed social schedule induced by the lockdown could allow them to wake up later without using an alarm clock and increase their sleep duration\textsuperscript{12}. In the older group, the opposite result was observed. An increase in sleep duration was documented in early chronotype, and no effect was observed in the late chronotypes. It is possible that this population could not perform some early social activity like group sports, but other explanations may also be possible. Unlike previous studies, we did not find an effect on social jetlag, even when we examined subpopulations: both MSW and MSF were delayed similarly in all subpopulations. Unlike previous studies, our control measurements were conducted between the first and second lockdowns, hence within a period of time that even without lockdown included many social changes (e.g., work from home, and distant education for some children age groups). Both misalignments of daily rhythms (as a result of shift work, frequent time-zone travel and even social jetlag) and short sleep duration have adverse effects on our immune system function\textsuperscript{32}. Studies showed that these factors increase infection probability\textsuperscript{28,30,36–38} and morbidity in response to viral infection\textsuperscript{29}, and reduce vaccine protection\textsuperscript{39,40}. Therefore, it is possible that the beneficial effects of lockdown in fighting COVID-19 results not only from the reduced social interactions (clearly observed in our study), but can also be attributed to the reduction in social jetlag and increase in sleep duration, which are beneficial to the immune system\textsuperscript{39,40}.

We found that in general the lockdown resulted in lower mood levels. This effect was evident in both age groups, but the decline was significant only in the early chronotypes of the younger group, which had higher mood levels before the lockdown and lower mood levels during the lockdown compared with older participants (Fig. 4A, B). A similar pattern was found in the number of steps per day, which decreased mainly in the younger participants. Such association between levels of physical activity and mood during the COVID-19 pandemic was previously reported\textsuperscript{41}. Interestingly, this group also did not increase sleep duration, in contrast with the young late chronotypes, which may contribute to the opposite effects of the restrictions on mood.

Stress levels did not change in our general study population, but interestingly that can be attributed to opposite effects in men and women: stress levels in men significantly decreased during the lockdown, but significantly increased in women (Fig. 3A). Women also experienced a more considerable decrease in the number of social encounters (Fig. 3B) — 40% vs. 20% in men. Physical activity and social interactions, including casual everyday meetings (e.g., with a seller at a coffee shop or a public transportation driver), are known to be correlated with mood, have beneficial effects for clinical or subclinical depression or anxiety, and as a means of upgrading life quality and well-being\textsuperscript{42–46}. Positive mood has been associated with enhanced immune function, while negative emotions such as stress, depression, and loneliness are correlated with the suppression of the immune system\textsuperscript{47–49}.

The opposite response of men and women in stress levels could result from social differences between them. In Israel, more women (21%) lost their workplace (fired or sent to unpaid vacation) compared to men (15%, The Israeli Ministry of Finance). Similarly, a study conducted in the US found that mothers with young children have reduced their work hours four times more than fathers\textsuperscript{50}. Moreover, schools and daycares were closed during lockdowns and parents had to stay at home with their children. Studies in Germany reported that during the COVID-19 lockdown men were more concerned about paid work while women were more worried about childcare and shouldered more childcare work\textsuperscript{51,52}. Furthermore, mothers experienced a greater decline in employment satisfaction while fathers’ well-being was less affected and their family satisfaction even increased\textsuperscript{53}. Another factor that could affect stress levels is domestic violence against women, which according to Israeli Police data, increased during lockdowns. Specifically, during March 1 and April 15 there was a 16% increase in the number of cases opened and 31% increase in phone calls to the police emergency line reporting domestic violence against women, compared to the same period last year (Israel Police data).

Together, our and others’ results\textsuperscript{54} suggest that young early chronotypes who did not increase their sleep duration, reduced activity level and suffered from significantly reduced mood, and women, who suffered an increase in stress levels and a greater
decline in social encounters, are more adversely affected from the lockdowns. These results are important by themselves, and suggest that these subpopulations, and especially early young chronotypes who usually are not in the focus of social support, should get special attention during lockdowns. Moreover, because these negative effects of the lockdowns can also suppress the immune system and result in higher infection probability and morbidity in response to viral infection and reduced vaccination efficiency, it is even more important to support the more negatively affected groups, and offer them guidance and recommendations on how to reduce these negative effects. We suggest that such knowledge can enable authorities to focus specific support actions to specific subpopulations to improve well-being, and consequently also the immune function.

The COVID-19 lockdown, together with our data collection framework, offer a unique opportunity to assess the effects of more relaxed social schedules on a multitude of daily rhythms and well-being indicators, which can be beneficial in our battle against the COVID-19 pandemic, and for the post COVID-19 “new normal” life. Using the combination of data collected from wearable devices and self-reported questionnaires from the same individuals, we are able to evaluate the effects of the lockdown on different ages, chronotypes and sexes for various well-being indicators. Based on the observed effects, we can suggest some recommendations for better coping with lockdowns and enhancing our immune function. These recommendations include proper sleep timing and duration, engagement in physical activity, and proper light exposure during the day and darkness during the night. The ability to work from home during the COVID-19 pandemic and the more flexible schedule may allow us to recommend and maintain healthier daily schedules to improve well-being and health, including reduced infection chances and severity, which can be implemented during, but even more importantly after, the COVID-19 pandemic.

Data availability

Because of the sensitive nature of some of the variables collected, the institutional review board (IRB)-approved protocol does not permit individual-level data to be made unrestricted and publicly available. Researchers interested in obtaining restricted, anonymized versions of this individual-level data should contact the authors to inquire about obtaining an IRB-approved institutional data sharing agreement.

Acknowledgements

This research was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 3409/19), within the Israel Precision Medicine Partnership program.

Competing interests

The authors declare no competing interests.

References


Appendix A: Subjects Recruitment and Engagement

In order to recruit subjects and keep them engaged throughout the study, we hired the services of a professional survey company. Potential participants were identified through advertisements in social media. The survey company was responsible to guarantee that participants met the study’s requirements, including their willingness to fill a daily questionnaire and wear a smartwatch during the entire study. Participants were met face-to-face and received detailed explanation about the study, after which they were requested to sign a consent form. Then, participants were asked to fill a one-time enrollment questionnaire and to install two apps on their smartphones: the Fitbit app which was used to collect data from their smartwatch, and the dedicated PerMed app which we developed.

The PerMed app was designed to passively collect smartphone sensors data, as well as self-reported daily questionnaires. The app also allows the participants to grant access to their smartwatch data, currently supporting Fitbit® and Garmin® devices. Data collected by the app (smartphone sensors data and self-reported questionnaires) as well as data collected from the smartwatch arrive to PerMed back-end servers and are stored securely.

Data in the PerMed pilot study were collected from four main sources: (1) the daily questionnaire, (2) smartphone sensors, (3) smartwatch, and (4) electronic medical records. In the current study we focused on data collected by the smartwatch and the answers to the daily questionnaires filled through the PerMed App.

In order to improve the quality and reliability of the data and to ensure its continuous collection, we applied the following measures. First, participants who did not fill the daily questionnaire by 7pm, received a notification in their mobile app to fill the questionnaire. Second, we developed a dedicated dashboard. The dashboard, which was monitored regularly, helped us identifying data collection issues, such as participants who did not fill the daily questionnaires or participants who did not wear their smartwatches. Such participants were contacted by the survey company, and were encouraged to cooperate better. The dashboard also helped us to identify issues that were not related to participants’ cooperation, such as bugs in the mobile app. This identification allowed us to respond faster and provide timely solutions.

To illustrate the quality of the collected data, Fig. 5 presents some statistics regarding the 169 participants included in the current study. Fig. 5A shows the distribution of the number of days each participant spent in our study. As can be seen from the figure, the vast majority of participants spent at least two months in the study. Fig. 5B presents the distribution of the daily questionnaire’s fill rate. As can be seen from the figure, the vast majority of participants filled the daily questionnaires in more than 60% of the days they spent in the study. Fig. 5C presents the distribution of coverage rate for smartwatch data. As can be seen from the figure, the vast majority of the participants had smartwatch data for more than 80% of the days they spent in the study.

Figure 5. Data quality: (A) Days since joining distribution, (B) Questionnaires fill rate, (C) Smartwatch collection rate. The X-axis represents the coverage ratio, and the Y-axis represents the ratio of the number of individuals with that coverage ratio or above.
Appendix B: The daily questionnaire

The daily questionnaire included the following eight questions:

1. How is your mood today?
   - Awful (-2)
   - Bad (-1)
   - OK (0)
   - Good (1)
   - Excellent (2)

2. How many hours of sleep did you have last night?

3. How would you define your last night sleep quality?
   - Awful (-2)
   - Bad (-1)
   - OK (0)
   - Good (1)
   - Excellent (2)

4. Try to remember how many minutes of sports activity you performed on the last day?

5. How would you describe the level of your stress during the last day?
   - Very Low (-2)
   - Low (-1)
   - Medium (0)
   - High (1)
   - Very high (2)

6. Try to estimate the number of people you have been with in the last day for up to 2 meters away?

7. Have you been diagnosed by a qualified medical professional with any of the following diseases? I was not diagnosed by a doctor with any disease
   - Flue
   - Covid-19
   - Strep throat
   - Cold
   - Other

8. Have you experienced one or more of the following symptoms in the last 24 hours?
   - My general feeling is good and I have no symptoms
• Heat measured above 37.5
• Cough
• Sore throat
• Runny nose
• Headache
• Shortness of breath
• Muscle aches
• Weakness / fatigue
• Diarrhea
• Nausea / vomiting
• Chills
• Confusion
• Loss of sense of taste / smell
• Another symptom
Appendix C: Detailed Descriptive Statistics

Table 3 presents information about the mean and number of participants for each of the 12 examined well-being indicators and each subpopulation, for the time period before the lockdown and for the time period during the lockdown.
### Table 3: Descriptive statistics: a breakdown by subpopulation. Each row represents a single well-being indicator, either before or during the lockdown period. Each column represents a subpopulation. Each entry contains the mean value and the number of individuals in that subpopulation in parenthesis.
Appendix D: Between-Subjects Effects

Table 4 the between-subjects effects for the Mixed ANOVA test that was conducted for each of the examined well-being indicators. Each row represents a single indicator (and thus also a single test) and each column represents the examined factor/interaction. Each entry represents the p-value of the corresponding effect, where statistically significant effects are marked with asterisks.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Age Group</th>
<th>Gender</th>
<th>Chronotype</th>
<th>Age Group Gender</th>
<th>Age Group Chronotype</th>
<th>Gender Chronotype</th>
<th>Age Group Chronotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mood</td>
<td>0.82</td>
<td>0.08*</td>
<td>0.03**</td>
<td>0.47</td>
<td>0.41</td>
<td>0.17</td>
<td>0.63</td>
</tr>
<tr>
<td>stress</td>
<td>0.10</td>
<td>0.00***</td>
<td>0.02**</td>
<td>0.79</td>
<td>0.76</td>
<td>0.07*</td>
<td>0.22</td>
</tr>
<tr>
<td>sleep quality</td>
<td>0.90</td>
<td>0.92</td>
<td>0.02**</td>
<td>0.95</td>
<td>0.95</td>
<td>0.24</td>
<td>0.35</td>
</tr>
<tr>
<td>sleep duration</td>
<td>0.16</td>
<td>0.31</td>
<td>0.02**</td>
<td>0.32</td>
<td>0.77</td>
<td>0.72</td>
<td>0.12</td>
</tr>
<tr>
<td>sport duration</td>
<td>0.03**</td>
<td>0.23</td>
<td>0.11</td>
<td>0.83</td>
<td>0.17</td>
<td>0.34</td>
<td>0.16</td>
</tr>
<tr>
<td>encounters</td>
<td>0.02**</td>
<td>0.71</td>
<td>0.10*</td>
<td>0.36</td>
<td>0.27</td>
<td>0.27</td>
<td>0.51</td>
</tr>
<tr>
<td>Smartwatch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>resting heart rate steps</td>
<td>0.04*</td>
<td>0.01***</td>
<td>0.57</td>
<td>0.79</td>
<td>0.76</td>
<td>0.69</td>
<td>0.02**</td>
</tr>
<tr>
<td>sleep duration</td>
<td>0.01***</td>
<td>0.19</td>
<td>0.11</td>
<td>0.60</td>
<td>0.02**</td>
<td>0.84</td>
<td>0.12</td>
</tr>
<tr>
<td>MSF</td>
<td>0.00***</td>
<td>0.88</td>
<td>0.00***</td>
<td>0.44</td>
<td>0.07*</td>
<td>0.24</td>
<td>0.36</td>
</tr>
<tr>
<td>MSW</td>
<td>0.00***</td>
<td>0.23</td>
<td>0.00***</td>
<td>0.43</td>
<td>0.61</td>
<td>0.23</td>
<td>0.46</td>
</tr>
<tr>
<td>SJL</td>
<td>0.00***</td>
<td>0.95</td>
<td>0.00***</td>
<td>0.43</td>
<td>0.02**</td>
<td>0.62</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 4. Between-subjects effects. Each row represents a Mixed ANOVA test for a single well-being indicator. Each column represents a single factor/interaction, and each entry represents the p-value of the corresponding effect. Statistically significant effects are marked with asterisks.

4.1 Age groups

Looking at age groups, it appears that the younger age group’s self-reported sport duration is lower than that of the older age group (an average number of 24.67 minutes vs. 33.09 minutes, p = 0.03, Cohen's d = 0.47). The steps count for the younger age group (as measured by the smartwatch) is higher than that of the older age group (an average number of 8826.05 steps vs. 7337.83 steps, p = 0.01, Cohen's d = 0.33). W also observe that the younger age group meet on average more people per day compared to the older age group (an average number of 11.31 encounters vs. 7.99 encounters, p = 0.02, Cohen's d = 0.42). It should be noted that the entire examined period is during the COVID-19 pandemic, and therefore older people tend to keep themselves more isolated to reduce risks. The results also show that the resting heart rate of the older age group is lower than that of the younger age group (an average hour of 4:05am vs. 3:26am, p = 0.04, Cohen's d = 0.30). As for sleep habits, younger people had later MSF and MSW and larger SJL than those of older people (an average hour of 4:33am vs. 3:42am, p < 0.01, Cohen's d = 1.23, an average hour of 4:05am vs. 3:26am, p < 0.01, Cohen's d = 1.01, an average of 0.93 hours vs. 0.67, p < 0.01, Cohen's d = 0.61). This is in-line with the literature where it is known that people’s chronotype becomes earlier with age. No significant differences were found in mood, stress, sleep quality, and sleep duration (both questionnaire and smartwatch).

4.2 Gender groups

Significant differences between men and women were found in mood and stress. It can be seen that men have a higher mood than women in general (an average level of 0.89 vs. 0.74, p = 0.08, Cohen's d = 0.24). Moreover, men are less stressed compared to women (an average level of -0.83 vs. -0.57, p < 0.01, Cohen's d = 0.42). We also observe that men’s resting heart rate is lower than that of women (an average of 60.65 bpm vs. 64.05 bpm, p = 0.01, Cohen's d = 0.61). These results seem to be consistent, as lower stress is likely to lead to better mood, and lower resting heart rate is known to be associated with lower stress levels. Interestingly, men sleep less (as measured by the smartwatch) than women (an average of 6.86 hours vs. 7.29 hours, p < 0.01, Cohen's d = 0.57). No significant differences were found in sleep quality, sleep duration - questionnaire, sport duration, encounters, steps, MSF, MSW and SJL.

4.3 Chronotype groups

Several significant differences were found between late chronotypes and early chronotypes. Late chronotypes have a better mood (an average level of 0.91 vs. 0.72, p = 0.03, Cohen's d = 0.29) and are less stressed than early chronotypes (an average
level of -0.80 vs. -0.06, \( p = 0.02, \text{Cohen's } d = 0.23 \). Furthermore, the sleep quality of late chronotypes is higher (an average level of 0.57 vs. 0.33, \( p = 0.02, \text{Cohen's } d = 0.37 \)) as well as their self-reported sleep duration (an average of 6.51 hours vs. 6.19 hours, \( p = 0.02, \text{Cohen's } d = 0.46 \)). As expected, late chronotypes had a later MSF (an average hour of 4:58am vs. 3:17am, \( p < 0.01, \text{Cohen's } d = 1.93 \)), a later MSW (an average hour of 4:28am vs. 3:03am, \( p < 0.01, \text{Cohen's } d = 1.55 \)) and a higher SJL (an average of 0.98 hours vs. 0.62, \( p < 0.01, \text{Cohen's } d = 0.74 \)). Late chronotypes also had a higher number of encounters per day compared to early chronotypes (an average number of 10.86 encounters vs. 8.44 encounters, \( p = 0.01, \text{Cohen's } d = 0.38 \)). No significant differences were found in sport duration, resting heart rate, steps and sleep duration recorded by the smartwatch.

### 4.4 Age Group & Chronotype Interaction

A significant interaction was found between age group and chronotype for the following indicators: sleep quality (\( p = 0.09 \)), steps (\( p = 0.02 \)), MSF (\( p = 0.07 \)) and SJL (\( p = 0.02 \)).

We observe that the sleep quality of the younger age group is highly influenced by chronotype (Fig. 6A). Specifically, the sleep quality of younger late chronotypes is higher than that of younger early chronotypes (an average level of 0.66 vs. 0.26). In the older age group, late chronotypes’ sleep quality is better than that of early chronotypes, but the difference was lower (an average level of 0.48 vs. 0.41). We can also see that for late chronotypes the sleep quality is better for the younger age group while in early chronotypes it is the opposite - the sleep quality of the older age group is better (Fig. 6A). Post-hoc analysis revealed that the change in sleep quality between the two chronotypes groups was significant in the younger age group (\( p < 0.01, \text{Cohen's } d = 0.67 \)) and not in the older age group.

As for the steps count, we observed that for the younger age group, the number of steps is higher for late chronotypes compared with early chronotypes (an average of 9031.08 steps vs. 8621.02). In the older group, it is the opposite, where the steps count for late chronotypes is lower than that of early chronotypes (an average of 6169.18 steps vs. 8506.47). It appears that for early chronotypes there is no such different between age groups, while for late chronotypes there is a relatively large difference of 2861.901 steps between the age groups (Fig. 6B). Post hoc analysis revealed that the change in steps between the two chronotypes groups was significant in the older age group (\( p < 0.01, \text{Cohen's } d = 0.68 \)) and not in the older age group.

As expected, the MSF value of late chronotypes is higher than that of early chronotypes for both age groups (young: an average value of 5:33am for late chronotype vs. 3:33am for early chronotypes, old: an average value of 4:24am for late chronotypes vs. vs. 3:11am for early chronotypes). However, the difference between chronotypes in the younger age group is higher than in the older group (Fig. 6C). Post-hoc analysis revealed that the change in MSF between the two chronotypes groups was significant in both the older age group (\( p < 0.01, \text{Cohen's } d = 1.87 \)) and in the younger age group (\( p < 0.01, \text{Cohen's } d = 2.02 \)).

It can also be seen that the SJL of late chronotypes changes with age. More specifically, the SJL of the younger age group is higher than the SJL of the older age group (an average of 1.22 hours vs. 0.74 hours). The difference between age groups for early chronotypes is small with an average SJL of 0.60 hours for the older group vs. 0.64 hours for the younger age group. In general, the SJL of late chronotypes is higher than that of early chronotypes in both younger and older age groups (Fig. 6D). Post-hoc analysis revealed that the change in SJL between the two chronotypes groups was significant in the younger age group (\( p < 0.01, \text{Cohen's } d = 0.85 \)) but not in the older age group.
Figure 6. Interactions between Age and Chronotype for: (A) sleep quality, (B) steps, (C) MSF and (D) SJL. The y-axis represents the Estimated Marginal Means (EMM) for the examined well-being indicator.

4.5 Gender & Chronotype Interaction

There was a significant interaction between gender and chronotype with regard to stress ($p = 0.07$).

As seen in Fig. 7A, in men, stress levels of late chronotypes is lower than that of early chronotypes (an average level of -1.01 vs. -0.66), while in women, the difference between late and early chronotype is much smaller (an average level of -0.59 vs. -0.55). Post-hoc analysis revealed that the change in stress levels between the two chronotype groups was significant in men ($p < 0.01$, Cohen's $d = 0.56$), but not for women.
Figure 7. Interactions between Gender and Chronotype for stress. The y-axis represents the Estimated Marginal Means (EMM) for the examined well-being indicator.

4.6 Age Group & Gender & Chronotype Interaction

We found a significant interaction between Age group, Gender and Chronotype in resting heart rate ($p = 0.02$) and sleep duration recorded by the smartwatch ($p = 0.05$). The interaction suggests that there are significant differences between various combinations of Age Group, Gender and Chronotype with respect to these two indicators (Fig. 8). Post-hoc analysis revealed that the change of resting heart rate between the two chronotypes groups was not significant. However, the change between men and women was significant in the older age group for early chronotypes ($p < 0.01$, $d = 0.75$) and in the younger age group for late chronotypes ($p < 0.01$, $d = 1.03$). The change of sleep duration between the two chronotypes groups was significant in the older age group for men ($p = 0.03$, $d = 0.57$) and the change between men and women was significant in the older age group for early chronotypes ($p < 0.01$, $d = 0.83$) and in the younger age group for late chronotypes ($p = 0.01$, $d = 0.65$).
**Figure 8.** Interactions between Age, Gender and Chronotype for: (A+B) resting heart rate, and (C+D) sleep duration recorded by the smartwatch. The y-axis represents the Estimated Marginal Means (EMM) for the examined well-being indicator.