Reliability of heart rate mobile apps in young healthy adults: exploratory study and research directions

Maria Parpinel, Laura Scherling and Stefano Lazzer
Department of Medicine, University of Udine, Udine, Italy

Vincenzo Della Mea
Department of Mathematics, Computer Science and Physics, University of Udine, Udine, Italy

ABSTRACT

Recently, a number of smartphone apps appeared that allow for heart rate measurements basing on the photoplethysmography principle. In fact, almost every smartphone now has a camera with flash that could be used for that. Some studies appeared on the reliability of some of those apps, with heterogeneous results. The present study aims at adding up evidence in particular during physical activity, by comparing three apps on two different platforms (iOS and Android), on a broad range of heart rates. As gold standard, heart rate has been measured with a traditional heart rate monitor.

The results suggest that heart rate apps might be used for measuring heart rate for fitness aims for many individuals, but further research is needed to i) analyse influence of smartphone features; ii) identify personal factors hindering measurements and iii) verify reliability on different measurement sites.

Keywords: exercise, fitness trackers, heart rate, mobile applications, physical fitness, telemedicine
INTRODUCTION

Recently, a number of mobile applications (apps) appeared that allow for heart rate measurements basing on the photoplethysmography principle\(^1\); they measure the changing absorbance of blood inside human tissue during the heartbeat cycle. The first implementation on smartphone, using its camera and flash, has been described by Pelegris et al.;\(^2\) however, a number of mobile applications are nowadays available, some of which are also described in scientific literature.\(^3,4\) Few independent studies also appeared that evaluated the effectiveness of such apps. For example, in children, it seems that they are not sufficiently precise.\(^5\) On healthy adults, results are more encouraging, possibly due to a better compliance and also finger size.\(^6,7\)

While some recent GPS watches exploit the reflectance pulse oximetry principle, their usual target is dedicated sports people. Those practising fitness or light/moderate sport activity might be interested in using a more generic device like the smartphone for casual readings of the heart rate, with some expected guarantee of reliability.

No study compared applications and devices on healthy adults during physical activity, and this drove the design of the present study. For the above-mentioned reason, the present paper aims at adding up evidence, by comparing three apps chosen among the most downloaded on two different smartphone platforms (iOSs and Android), and on a broad range of heart rates measured on healthy adult subjects during physical activity. The study was aimed at identifying some possible research questions to be answered in a further extension.

METHODS

Subjects

Ten volunteer young adults (four males and six females) participated in the study, and we collected their anthropometric and physical activity data. We conducted the study in agreement with the declaration of Helsinki and collected informed consent from participating subjects. However, according to our institution regulations, it did not need ethical approval by our ethics committee.

Devices

We used two smartphones representative of iOSs and Android operating systems: iPhone 5 (with iOS 7.1.2) and Huawei Ascend G700 (with Android 4.2.1). Gold standard was provided by a belt heart rate monitor (Polar RS300X), with precision ±1% (minimum ±1 bpm) and valid range between 15 and 240 bpm.

Applications

Applications have been chosen among those listed in the health and fitness category of the Google Play store for Android devices. We identified the most downloaded ones\(^7\) and then selected those available for iPhone too. We remained with three applications: Instant Heart Rate (INR), Cardiograph (CAR) and Heart Beat Rate (HBR). These are to be considered only a sample of applications, useful to understand possible issues.

Study design and analysis

Heart rate has been measured at rest and after physical activity with a laboratory cycle ergometer (Daum Electronic Ego Bike Cardio-Pro). To collect a wide range of heart rates, exercise sessions were organised in three 15-min workouts at increasing intensity (90, 110 and 130 W), repeated for three days, with resting time between workouts. Heart rate measurements occurred every 5 min, with two smartphones at a time on left and right hand.

Average difference for each individual subject and overall were calculated between app values and corresponding gold standard measurements and compared using the Spearman correlation coefficient.

RESULTS

The total number of heart rate measurements made was 1080, by combining rest, activity intensity, device and app and repeating for three days. This means that each app was evaluated on 180 measurements on each platform. Twelve measurements were discarded for problems in reading. The heart rate as measured by the gold standard heart rate monitor ranged from 53 to 192 bpm and from 49 to 193 bpm by apps.

Figure 1 shows an overall chart of the correspondence between gold standard and the three apps, executed on both platforms.

Average difference between gold standard and applications on both devices and corresponding Spearman correlation coefficients are shown in Table 1. All correlations appear high and statistically significant (\(p<0.0001\)). However, as it can be seen, all applications underestimate heart rate, with some differences among them and between devices.

One further step was to study individual influence on results. For this, we calculated the average difference for each subject by app and device where the two largest differences for each column are highlighted in bold. Subject 4 systematically had largest differences between app measured heart rate and gold standard. We found subject 4 being a female, with the least height, weight and BMI, and after physical activity, done by healthy individuals that are possibly interested in fitness.

DISCUSSION

In this preliminary evaluation of reliability of heart rate measurements by means of mobile apps, we focused on one possible field of application, i.e. self-measurements during or after physical activity, done by healthy individuals that are possibly interested in fitness.

The heart rate measured by the tested apps well correlated with the gold standard, although it is mostly underestimated. This result is in line with other comparative works on adults.\(^3,6\) However, not all apps have disclosed equal behaviour: one of the tested apps shown decreased reliability at higher frequencies, to be further investigated. Issues at high
frequencies were already reported, and could be possibly linked to a couple of causes. First of all, the camera module of smartphones could allow for variable frame rates in the acquisition of the images needed for heart rate measurement. Secondly, the computation carried out by apps may rely on algorithms more or less computationally intensive, thus running at different speeds on differently powerful smartphones.

We also preliminary investigated individual differences. While some app instructions state that cold fingers may give difficulties in measurements, no other possible issues are mentioned. However, in one case, we found that possibly hand size or finger size could have influenced measurement quality, and this is an indication for further experiments to be done to identify subpopulations for which heart rate measurement is systematically less reliable. In the work of Ho et al., the apps were tested not only on the finger, but also on the earlobe: this could be a workaround for the mentioned problems, but needs further validation.

Finally, while we tested two different devices without finding significant differences, it is possible that some hardware features may influence the quality of measurement: for example, camera resolution, LED light features, processor speed or even smartphone size. In fact, Pelegris et al. mentioned that one of their two implementations processed much less frames per second, thus limiting the maximum heart rate that could be possibly measured. These factors could be more crucial for entry-level smartphones, which are less powerful and thus slower in computation, and this could be an issue for a quick execution of the algorithms normally adopted for photoplethysmography implementation (weighted moving average, fast Fourier transform and time-frequency analysis). Also, the position of the camera (centred or in a corner) might interact with smartphone width in making the finger position more or less comfortable and positioning more or less precise. Ideally, a large smartphone might be more usable if the camera is centred, while a smaller one could be better suitable if camera is in a corner, but this needs to be confirmed by experimentation.

Thus, we need to identify when they are not and why, and for this, we envisage the following research directions:

- study of influence of smartphone features on heart rate measurement quality, by evaluating more devices with heterogeneous characteristics;
- better understanding of personal factors limiting heart rate measurements by collecting more anthropometric data (including hand size and microcirculation features) on a larger subject group;
- testing different measuring sites – for example, on the arm- where smartphone bands are commonly fit during physical activity.

Only after that, smartphone-measured heart rate could be eventually considered usable for non-medical applications.

Table 1 App versus gold standard average differences by subject, smartphone and app and overall correlations

<table>
<thead>
<tr>
<th></th>
<th>Android</th>
<th>iPhone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1.39</td>
<td>-1.61</td>
</tr>
<tr>
<td>2</td>
<td>-15.56</td>
<td>-7.11</td>
</tr>
<tr>
<td>3</td>
<td>-17.39</td>
<td>-8.22</td>
</tr>
<tr>
<td>4</td>
<td>-20.61</td>
<td>-12.89</td>
</tr>
<tr>
<td>5</td>
<td>-2.61</td>
<td>-2.17</td>
</tr>
<tr>
<td>6</td>
<td>1.17</td>
<td>-1.28</td>
</tr>
<tr>
<td>7</td>
<td>-0.11</td>
<td>2.61</td>
</tr>
<tr>
<td>8</td>
<td>-0.67</td>
<td>-1.78</td>
</tr>
<tr>
<td>9</td>
<td>-21.50</td>
<td>-12.72</td>
</tr>
<tr>
<td>10</td>
<td>-2.56</td>
<td>-1.89</td>
</tr>
<tr>
<td>Average</td>
<td>-8.122</td>
<td>-3.61</td>
</tr>
<tr>
<td>Spearman correlation coefficient</td>
<td>0.844</td>
<td>0.981</td>
</tr>
</tbody>
</table>
REFERENCES


