Contents

Original Papers

Motivations and Key Features for a Wearable Device for Continuous Monitoring of Breathing: A Web-Based Survey (e1)
Joseph Prinable, Juliet Foster, Alistair McEwan, Paul Young, Euan Tovey, Cindy Thamrin. ................................................................. 2

Measurement of Skin Induration Size Using Smartphone Images and Photogrammetric Reconstruction: Pilot Study (e3)
Ronald Dendere, Tinashe Mutsvangwa, Rene Goliath, Molebogeng Rangaka, Ibrahim Abubakar, Tania Douglas. .............................. 14

Heart Rate Monitoring Apps: Information for Engineers and Researchers About the New European Medical Devices Regulation 2017/745 (e2)
Michael Lang. .................................................................................................................. 21
Motivations and Key Features for a Wearable Device for Continuous Monitoring of Breathing: A Web-Based Survey

Joseph Barry Prinable1,2, B Eng (Hons); Juliet M Foster2, PhD; Alistair L McEwan1, BE, PhD; Paul M Young2, PhD; Euan Tovey2, PhD; Cindy Thamrin2, PhD

1Faculty of Engineering and Information Technologies, School of Electrical and Information Engineering, University of Sydney, Sydney, Australia
2Woolcock Institute of Medical Research, University of Sydney, Glebe, Australia

Corresponding Author:
Joseph Barry Prinable, B Eng (Hons)
Faculty of Engineering and Information Technologies
School of Electrical and Information Engineering
University of Sydney
Rm 340, J03
Sydney, 2006
Australia
Phone: 61 2 9351 7256
Email: jpr6895@uni.sydney.edu.au

Abstract

Background: Analysis of patterns of breathing over time may provide novel information on respiratory function and dysfunction. Devices that continuously record and analyze breathing rates may provide new options for the management of respiratory diseases. However, there is a lack of information about design characteristics that would make such devices user-friendly and suitable for this purpose.

Objective: Our aim was to determine key device attributes and user requirements for a wearable device to be used for long-term monitoring of breathing.

Methods: An online survey was conducted between June and July 2016. Participants were predominantly recruited via the Woolcock Institute of Medical Research database of volunteers, as well as staff and students. Information regarding the survey, a consent form, and a link to a Web-based questionnaire were sent to participants via email. All participants received an identical survey; those with doctor-diagnosed asthma completed an extra questionnaire on asthma control (Asthma Control Test). Survey responses were examined as a group using descriptive statistics. Responses were compared between those with and without asthma using the chi-square test.

Results: The survey was completed by 134 participants (males: 39%, median age group: 50-59 years, asthma: 57%). Of those who completed the Asthma Control Test, 61% (47/77) had suboptimal asthma control. Of the 134 participants, 61.9% (83/134) would be willing to wear a device to monitor their breathing, in contrast to 6.7% (9/134) who would not. The remaining 31.3% (42/134) stated that their willingness depended on specific factors. Participants with asthma most commonly cited their asthma as motivation for using a wearable; the most common motivation for use in those without asthma was curiosity. More than 90% of total participants would use the device during the day, night, or both day and night. Design preferences among all users included a wrist watch (nominated by 92.5% [124/134] for both day and night use, out of four body sites), the ability to synchronize breathing data with a mobile phone or tablet (81.3%, 109/134), overnight power charging (33.6%, 45/134), and a cost of ≤Aus $100 (53.7%, 72/134).

Conclusions: We have explored the motivations and likelihood for adopting wearable technologies for the purpose of monitoring breathing and identified user preferences for key design features. We found participants were motivated to adopt a wearable breathing monitor irrespective of health status, though rationale for use differed between those with and without asthma. These findings will help inform the design of a user-acceptable wearable device that will facilitate its eventual uptake in both healthy and asthma populations.

(JMIR Biomed Eng 2017;2(1):e1) doi:10.2196/biomedeng.7143
KEYWORDS
asthma; wearable; breathing; e-health

Introduction

Asthma is a serious public health problem affecting over 300 million people globally. Management challenges include the early prediction or warning of asthma attacks and optimizing the pharmaceutical management of the disease.

Monitoring of lung function over time is a widely accepted component of the assessment of asthma, both in clinical management of the disease as well as in research trials [1]. Some studies suggest it may also yield insights into the pathology of respiratory diseases and predict future risk of exacerbations [2-4]. In asthma, monitoring is usually based on standard lung function testing involving forced breathing maneuvers assessed periodically in a specialized respiratory laboratory, or by peak expiratory flow measured in a general practice and then in the patient’s home either daily or during periods of worsening symptoms. There is a paucity of research on continuous, real-time monitoring of breathing for general health or for management of asthma or other chronic diseases. This may be due in part to the lack of commercial technology to enable such monitoring in a manner that would be acceptable to users. One study has shown that monitoring respiratory rate could help predict the onset of exacerbations in chronic obstructive pulmonary disease [5]. However, it is not known whether monitoring of breathing could aid diagnosis or monitoring of asthma. Breathing monitoring may also provide rapid feedback to a patient during physical exertion or breathing exercises during exacerbation episodes.

Several studies have investigated desirable features for a wearable device for health monitoring, from both a technical [6] and human-centered [7-11] perspective. These studies have provided guidelines on wearable design [7-9] and determined that user acceptability was dependent on factors such as fundamental needs/demonstrated benefit, enjoyment, and social value [10,11]. However, none of these studies sought to specifically determine the desired features for a wearable device used for long-term respiratory monitoring. At present, there are several modalities and locations on the body identified for respiratory monitoring: the ear, throat region, finger, wrist, and chest [12-14]. In the design and development of a device for this purpose, it is important to first identify, understand, and consider user preferences to increase user acceptance, satisfaction, and engagement [8,15].

The purpose of this study is to (1) explore the reasons why participants with or without asthma would potentially adopt new technologies to monitor breathing over time, and (2) evaluate device-specific attributes that would meet the expectation of users within these two groups. We chose to additionally study healthy individuals, not only as a basis for comparison with those with asthma to identify those preferences that are specific to asthma, but also due to the increasing interest in personal health monitoring in the general population as evidenced by the uptake of wearable devices that measure activity and other physiological life signs.

Methods

Study Design and Overview

An online survey was conducted between June and July 2016. A link to the survey was sent electronically to a subset (n=569) of the Woolcock Institute of Medical Research Volunteers Database based on the availability of a valid email address on record, as well as to staff and students at the Woolcock Institute. During the recruitment period, two rounds of recruitment emails were sent to the two lists, followed by a subsequent reminder email for each round. The Volunteers Database consists of members of the public who have previously given consent to be contacted about participation in research. The database comprises both healthy individuals (n=256) as well as those with asthma (n=1173). The exact number reached may differ due to constant additions or withdrawals from the database and the possibility of family members sharing a common email address. Inclusion criteria were (1) provision of informed consent, (2) completion of all responses, (3) no respiratory illness reported (for the healthy group), and (4) self-reported doctor’s diagnosis of asthma (for the asthma group). No incentives were offered for participation. The protocol for this study was approved by Northern Sydney Local Health District Human Research Ethics Committee (ethics approval #LNR/16/HAWKE99).

Survey

After clicking on the link to the survey, participants who provided informed consent proceeded to fill out an online questionnaire (see Multimedia Appendix 1) that took approximately 10-15 minutes to complete. The survey was designed to assess participant’s current use of technology, to explore their readiness to use a wearable, and to understand their attitude toward the potential usefulness of wearable technologies for monitoring breathing. Specifically, the survey aimed to identify usage preferences (eg, how long the user wishes to wear the device during the night/day) and feature preferences, such as the device form factor (eg, band, sticky patch, earpiece), body location (eg, wearable for neck, chest, ear, wrist), display, charge time, and price.

The survey also included demographic questions such as age, gender, educational and socioeconomic status, and doctor-diagnosed health conditions. Those who reported having a doctor diagnosis of asthma completed the Asthma Control Test (ACT) [16], a well-validated scale [17], which comprises five questions that assess asthma symptoms, use of medication, and the effect of asthma on daily functioning to determine overall asthma control status. The total score ranges from 5 (poor control of asthma) to 25 (complete control of asthma); a score of ≤19 indicates suboptimal control.

Statistical Analyses

Participant demographics were summarized using descriptive statistics. Results were compared between participants with self-reported doctor-diagnosed asthma versus those without doctor’s diagnosis of asthma.
asthma, using $t$ tests or Wilcoxon signed rank sum test depending on whether the data were normally distributed. Participants who were “unsure” of their asthma status were grouped with those participants without asthma. Questionnaire responses were compared between asthma and no asthma, between gender, and between age groups using chi-square tests. Statistical analyses were performed using SPSS v. 23 (IBM Corp.), and graphs were generated using Prism v. 7 (GraphPad Software Inc.).

**Results**

**Demographics**

In total, 156 participants responded but 2 did not provide informed consent and 20 failed to complete more than 50% of the survey and were omitted from analysis. Of the 134 participants who completed the survey (ie, 85.9% completion rate), 131 provided demographic information as shown in Table 1. Just under a third (29.1%, 39/134) of participants were male, and nearly two-thirds (60.0%, 79/134) had a university education. More than 10 participants were obtained in each age group. The average time to complete the survey was 13 minutes.

A total of 61.2% (76/134) participants reported doctor-diagnosed asthma: mean (SD) ACT score was 17.4 (5.2). Nearly two-thirds (62%, 47/76) of these had suboptimal asthma control based on the ACT.

**Technology and Device Use**

Participants demonstrated a high level of technology use: 88.8% (119/134) used a smart phone, 29.9% (40/134) used health monitoring devices such as a Fitbit, and a small percentage of participants used smart watches (5.2%, 7/134). Nearly two-thirds (59.7%, 80/134) used only one form of technology, 26.9% (36/134) used two forms of technology, and 3.0% (4/134) used three or more forms of technology. Examples of other specific technology or gadgets used were fitness trackers (11.9%, 16/134), tablet computers (11.9%, 16/134), music players (3.0%, 4/134), conventional mobile telephones (1.4%, 2/134), and electronic books (1.4%, 2/134). Only 8 participants (5.9%, 8/134) used no “other forms of technology or electronic gadgets”. Levels of technology use were similar in those with and without asthma.

**Motivation for Wearable Use**

Nearly two-thirds (61.9%, 83/134) of the total participants indicated that they would be willing to wear a device to monitor their breathing, 7.4% (10/134) would not, and the remaining 30.5% (41/134) stated that their willingness depended on specific factors, described later in this section. There were no significant differences in willingness to adopt a wearable device for monitoring breathing between the 40 participants who currently used health monitoring devices and the 94 who did not ($P=0.265$). Participants with asthma were more willing to wear a device to monitor their breathing, compared to those without asthma: 70% (54/77) versus 51% (29/57), $P=0.071$.

Regardless of whether or not they were willing to use a wearable, participants were asked to indicate one or more factors that would make them consider using a wearable. These are detailed in Figure 1. Out of all participants, more people who did not have asthma indicated “curiosity” (23%, 13/57 vs 10%, 7/77; $P=0.028$) or “I would like to track my performance during exercise” (30%, 17/57 vs 10%, 8/77; $P=0.004$) as a motivating factor to wear the device than those with asthma.

Females were more likely to use the device to track breathing patterns during stress and meditation compared to men (16%, 15/92 vs 3%, 1/39; $P=0.003$). Females were also more likely to use the device when they get breathless (9%, 8/92 vs 5%, 2/39; $P=0.002$) or if they had a known respiratory disease other than asthma compared to men (8%, 7/92 vs 0%, 0/39; $P=0.031$).

The ability to track breathing patterns during stress and meditation was a more common rationale for device use in younger than older age groups: 18-39 (37%, 7/19), 30-39, (4%, 1/27), 40-49 (13%, 2/15), 50-59 (19%, 5/26), 60-69 (3%, 1/31), older than 70 (0%, 0/13); $P=0.003$. Curiosity was also a more common rationale for use in younger people: 18-39 (42%, 8/19), 30-39 (22%, 6/27), 40-49 (7%, 1/15), 50-59 (4%, 1/26), 60-69 (6%, 2/31), older than 70 (8%, 7/92 vs 0%, 0/39; $P=0.031$).

A larger proportion of the 40 participants who already used a health monitoring device would wear one to monitor their breathing for their asthma or to track patterns during stress (48%, 19/40 for both), compared to those out of the 94 who did not currently use a device (29%, 27/94 for both; $P=0.036$).

Participants were asked to indicate whether any respiratory illnesses other than asthma were part of their motivation to wear a wearable. Only 8 reported that this was a motivating factor.
Table 1. Participant demographic information for the wearable survey study, stratified by health status.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total, n (%) (n=131)</th>
<th>No asthma, n (%) (n=55)</th>
<th>Asthma, n (%) (n=76)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender: Male</strong></td>
<td>39 (29)</td>
<td>22 (40)</td>
<td>17 (22)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-29</td>
<td>19 (15)</td>
<td>14 (25)</td>
<td>5 (7)</td>
</tr>
<tr>
<td>30-39</td>
<td>27 (20)</td>
<td>13 (23)</td>
<td>14 (18)</td>
</tr>
<tr>
<td>40-49</td>
<td>15 (11)</td>
<td>8 (15)</td>
<td>7 (9)</td>
</tr>
<tr>
<td>50-59</td>
<td>26 (20)</td>
<td>6 (11)</td>
<td>20 (26)</td>
</tr>
<tr>
<td>60-69</td>
<td>31 (24)</td>
<td>6 (11)</td>
<td>25 (33)</td>
</tr>
<tr>
<td>70+</td>
<td>13 (10)</td>
<td>8 (15)</td>
<td>5 (7)</td>
</tr>
<tr>
<td><strong>ACT, mean (SD)</strong></td>
<td>—</td>
<td>—</td>
<td>17.4 (5.2)</td>
</tr>
<tr>
<td><strong>Highest level of education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary school</td>
<td>21 (16)</td>
<td>6 (11)</td>
<td>15 (20)</td>
</tr>
<tr>
<td>Higher certificate or diploma</td>
<td>30 (23)</td>
<td>9 (16)</td>
<td>21 (28)</td>
</tr>
<tr>
<td>Bachelor degree or higher</td>
<td>79 (60)</td>
<td>39 (71)</td>
<td>40 (52)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1 (1)</td>
<td>1 (2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Socioeconomic status: Low SES</strong></td>
<td>25 (19)</td>
<td>9 (16)</td>
<td>16 (32)</td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment, full or part time</td>
<td>80 (63)</td>
<td>36 (65)</td>
<td>44 (61)</td>
</tr>
<tr>
<td>Employment, casual</td>
<td>12 (10)</td>
<td>7 (13)</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Currently unemployed</td>
<td>34 (27)</td>
<td>11 (20)</td>
<td>23 (32)</td>
</tr>
<tr>
<td><strong>Household income (Aus $)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$26,000</td>
<td>12 (9)</td>
<td>4 (7)</td>
<td>8 (11)</td>
</tr>
<tr>
<td>$26,000-$51,999</td>
<td>20 (15)</td>
<td>8 (15)</td>
<td>12 (16)</td>
</tr>
<tr>
<td>$52,000-$72,799</td>
<td>19 (14)</td>
<td>9 (16)</td>
<td>10 (13)</td>
</tr>
<tr>
<td>$72,800-$103,999</td>
<td>21 (16)</td>
<td>6 (11)</td>
<td>15 (20)</td>
</tr>
<tr>
<td>$104,000-$155,999</td>
<td>9 (7)</td>
<td>3 (5)</td>
<td>6 (8)</td>
</tr>
<tr>
<td>≥$156,000</td>
<td>23 (18)</td>
<td>15 (27)</td>
<td>8 (11)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>27 (21)</td>
<td>10 (18)</td>
<td>17 (21)</td>
</tr>
<tr>
<td><strong>Language other than English spoken at home</strong></td>
<td>19 (15)</td>
<td>11 (20)</td>
<td>8 (11)</td>
</tr>
</tbody>
</table>

---

*a131/134 participants who completed a survey provided demographic data.

*b55/57 participants who did not have doctor-diagnosed asthma provided demographic data.

*c76/77 participants who had doctor-diagnosed asthma provided demographic data.

*dA score of ≤19 indicates suboptimal asthma control.

*e1/131 participants who provided demographic data did not report their education status.

*fSocially disadvantaged at patient’s home address: “Disadvantaged” Socio-Economic Indexes For Area (SEIFA) quintile <3, “Advantaged” SEIFA quintile: 4-5 [18].

*g2/131 participants who provided demographic data did not provide employment information; “Currently unemployed” includes unpaid or volunteer work, engagement in home duties, or not being in the labor force.

*h27/131 participants who provided demographic data did not provide household income information.
**Those Willing to Use a Wearable Device**

When we restricted our analyses to the subgroup of those willing to use a wearable device only (61.9%, 83/134), the most common motivating factor to wear a device for those without asthma was “curiosity” (59%, 17/29; \( P = .026 \)). The most common motivating factor for people with asthma was “I have asthma” (83%, 45/54; \( P < .001 \)). No significant differences were observed between those with and without asthma in the other provided reasons. Figure 1 shows user motivation across this subgroup, stratified by self-reported, doctor-diagnosed asthma status, with participants able to select multiple responses.

**Those Who Would Not Use a Device**

In this subgroup (6.7%, 9/134), those without asthma stated they would not wear a device because they did not understand why monitoring breathing was important (eg, “I can’t see a reason why I would want to monitor my breathing”).

The reasons for not using the device in the four participants with asthma were that they felt their asthma was under control (eg, “Asthma is under control,” “I don’t get bad asthma attacks, just slight, not worth the bother”), or due to travel or cost (“I am overseas at this time,” “Such devices are too expensive”).

**Those Whose Willingness Depended on Specific Factors**

In this subgroup (31.3%, 42/134), 19 had asthma and 23 did not. The most common motivating factor for wearing a device

---

Figure 1. User motivation for those willing to use a wearable device, stratified by self-reported, doctor-diagnosed asthma status.
in people with asthma was “I have asthma” (83%, 14/19; \(P \leq .001\)). No significant factors were found for those without asthma in this subgroup.

Factors Affecting Wearable Use

The factors affecting wearability mentioned across all participants included design issues and user perception issues. In terms of design, the physical size, location, weight, and bulk of the device were common concerns. Related to these were user perception issues, such as comfort and inhibition of movement, discreetness, and how the device would be fitted to the body. Example of factors provided were “how comfortable and discrete the device is,” “how it’s worn,” “size would it inhibit normal movements and is it 24/7?”

Unappealing Factors

All participants were asked to select which factors would cause them to consider a wearable device unappealing (Figure 2). Of note, 26% (15/57) of participants without asthma did not see the usefulness of the device, compared to 9% (7/77) of those with asthma (\(P = .008\)). More participants without asthma would use a device to monitor breathing only if they were told to by a medical professional compared to those with asthma (39%, 20/49 vs 17%, 13/77; \(P = .005\)).

Figure 2. Unappealing factors for wearing a device, stratified by self-reported, doctor-diagnosed asthma.
Device-Specific Features

The device-specific features were themed into five different categories: wearability, cost, power features, display, and data synchronization. All 134 survey participants completed this section. In general, there were no differences between those who were current users of health monitoring devices and those who were not, in preference for form factor, length of usage, cost, display or data storage time preferences, unless otherwise indicated below.

Wearability

A majority (94.0%, 126/134) of respondents (with or without asthma) would use a wearable device during the day, night, or both day and night. Most users preferred to wear the device 5 nights/days a week or more (Figure 3). However, more out of those who already used a health monitoring device indicated they would use the device for 5 days or more a week (83%, 33/40), compared to those who did not already use a monitoring device (60%, 56/94; *P*=.01).

Furthermore, those with asthma said they would wear the device more often than those without asthma during both the night and day: 82% (63/77) with asthma versus 46% (26/57) without asthma would wear the device 5 days or more a week; *P*<.001. Those without asthma were also more likely to wear the device only during training: 26% (15/57) versus 5% (4/77); *P*=.001. No significance differences were found between health status and form factor for daytime use.

Frequency of daytime and nighttime use was higher in older people. For example, older participants predicted they were more likely to wear the device 5 days a week or more during the night: 18-39 (37%, 7/19), 30-39, (59%, 16/27), 40-49 (53%, 8/15), 50-59 (81%, 21/26), 60-69 (74%, 23/31), older than 70 (77%, 10/13); *P*=.026. Younger age groups were more likely to use the device during exercise than older age groups: 18-39 (47%, 9/19), 30-39 (15%, 4/27), 40-49 (7%, 1/15), 50-59 (4%, 1/26), 60-69 (10%, 3/31), older than 70 (8%, 1/13); *P*<.001.

There was a clear preference for a wrist band over other formats such as earbuds, and preferences were similar for day versus nighttime use (Figure 4). Men were more likely to wear a chest band during the day (38%, 15/39 vs 20%, 18/92; *P*=.043) compared to women. At night, men were also more likely to wear an earbud in the ear (28%, 11/39 vs 16%, 15/92; *P*=.044) but less likely to wear a wrist band (90%, 35/39 vs 97%, 89/92; *P*=.039) compared to women.

**Figure 3.** Total participant preference for how often the device is to be worn, separated by day and night use.

**Figure 4.** Total participant preference for the form of the device to be worn, separated by day and night use.
**Cost**
Over half (53.7%, 72/134) of the total participants would be happy to pay up to Aus $100 for a wearable respiratory monitor, 20.8% (28/134) would pay over Aus $100, and the remaining 25.3% (34/134) would use it only “if it were free.” No statistically significant differences were observed in responses by health status, different household income, age, or gender.

**Power Features**
The most popular waiting time for the device to charge was overnight (45/134, 33.6%) as opposed to within 2 hours (22.3%, 30/134), 1 hour (23.1%, 31/134), 30 minutes (11.1%, 15/134), or other (10.4%, 14/134). Charging time did not appear to be a critical factor in user preferences, with other responses provided as: “As long as it takes. Good if the recharging was no more than 2 hours” or “However long it took to charge.” No differences were observed between those with or without asthma.

**Display**
Participants selected between the three different displays shown in Figure 5, representing different formats to display current and past breathing data. No preference was found between display type (numerical information, 48/134; bar graph, 39/134; line graph, 47/134). There was no difference in display preference between those with asthma and without asthma, or between different age or gender groups.

Participants indicated that they would like to receive alerts when their breathing was problematic. Alerts were more popular in those with asthma than those without asthma: 79% (61/77) versus 63% (36/57); P=.048.

**Syncing and Data Storage**
The majority of participants (79.8%, 107/134) reported wanting to sync the device to their phone/tablet. The proportion was higher among those who already use a monitoring device (93%, 37/40). Less than half (45.5%, 61/134) wanted to sync the device with their computer. Those who selected “other” responded with “remote analysis and syncing with my GPs office,” “sync with sleep study,” or “cloud service.” Younger participants were more likely to report wanting to sync their breathing data (number of breaths per minute) with a phone or tablet than older participants: 18-39 (100%, 19/19), 30-39 (100%, 27/27), 40-49 (87%, 13/15), 50-59 (77%, 20/26), 60-69 (68%, 21/31), older than 70 (54%, 7/13); P=.001.

The majority of participants reported wanting to save their data for at least 1 week (58.9%, 79/134).

**Discussion**

**Principal Findings**
In this survey, we identified a number of reasons to adopt new technologies to monitor breathing in participants with or without asthma. In participants without asthma, the main factor that influenced motivation for using a wearable was curiosity. The ability to track breathing patterns during stress or meditation and fitness tracking were motivational factors for younger participants. In asthma, the main motivations for use were “having asthma” and the ability to track breathing patterns during periods of breathlessness. We found that most users were willing to wear the device continuously both day and night and that the most preferred device format was a wrist band, regardless of health status. Other desired features were alerts when breathing is problematic (for both asthma and non-asthma groups), the ability to synchronize data with a phone or tablet, a recharging period of every 24 hours, and cost of ≤ Aus $100.

**Motivation for Wearable Use**
Previous studies have found that perceived value has a significant influence on both potential and actual customers, with perceived value as an important factor influencing the consumer’s decision to adopt new products or services [7,11]. One of the most influential factors for people without asthma was curiosity, a factor that in previous research has been thought to increase initial interest and subsequent user engagement [19].

As might be anticipated, motivation for using a wearable device in asthma was different to those without asthma. In people with asthma, there appeared to be a desire to use breathing monitoring to gain greater control over the management of their asthma, particularly during episodes of breathlessness. An episode of extreme breathlessness during a respiratory exacerbation is often extremely frightening to both patients and their family members [20]. Provided that there has been sufficient testing and development of safe and reliable markers, detailed self-tracking breathing metrics could potentially help provide patients with an objective identifier or predictor of such episodes. This is especially important given that self-perception of airway narrowing is known to be poorer during an asthma exacerbation than at other times [20]. For family members, real-time monitoring may allow them to assist in supporting their relative with asthma in identifying symptom worsening and deciding when to seek emergency care, alongside traditional indicators.
Patients are known to employ a number of strategies to cope with breathlessness episodes, including breathing techniques and reduction of physical exertion [21]. A simple wearable device to measure breathing may provide objective monitoring and feedback during use of breathing techniques, and with the guidance of a health professional, has the potential to support patients to increase their physical activity in a safe manner. A monitor that directly and continuously measures breathing might provide a unique capability for immediate feedback that may not be achieved with currently available devices, such as those measuring wheezing sounds, peak flow, or lung mechanics. There are precedents for monitoring and feedback in asthma, for example, monitoring and feedback of medication use is acceptable and has been shown to increase medication use in adults and children [22,23].

The observed difference in the rationales for using a breathing monitoring device between participants with and without asthma indicates the need to collect separate data on the motivation for use and the utility and feasibility of wearables (for breathing or other purposes), in people with and without (different) health conditions. Conversely, the rationale for choosing not to adopt a wearable device for breathing monitoring was similar between those with and without asthma. The main reason given was a lack of perceived purpose or need for such a device, for example, because asthma was already “under control.” Indeed, there is a lack of direct evidence showing that the ambulatory monitoring of breathing patterns over time is useful for asthma. This is despite the disease being characterized by shortness of breath. However, indirect support comes from measurements made using breathing-based lung function tests [24], recent developments in the monitoring of wheeze [25], and data showing breathing patterns predictive of chronic obstructive pulmonary disease exacerbations [5]. The availability of a suitable wearable will enable further work showing utility in asthma management.

User Preference for Device Features

To the best of our knowledge, this is the first time user preferences for a wearable device aimed at respiratory health monitoring have been investigated. This is important as desired design features often come at a technical cost. The results of this study inform us which features are of high value and which features could be compromised in exchange for technical tradeoffs. Furthermore, acceptance of a new technology may be affected by the perceived risk or inconvenience posed by the device. Previous research suggests that factors such as wearability design, physical size, location, weight, and bulk may negatively impact perceived device value. Costly and complicated recording devices may result in low compliance [10].

There are little data available to suggest what constitutes acceptable levels for these features and for human factors in a breathing monitor wearable. In this study, we found significant user perception issues around comfort and inhibition of movement, discreetness, and where the device fits on the body. Our study also revealed that more than 90% of participants would wear the device both day and night, and more than 90% preferred a wrist-worn device. Comfort and frequency of use are likely to interact, with more comfortable devices used for longer.

Most users preferred a wrist band over other formats for site of monitoring; however, this may have been influenced by the type of devices most commonly available on the market at the time. We note that chest bands and ear buds were also identified as next preferred formats for monitoring and may have been selected by participants with existing exposure. Device design choice needs to be made in terms of both user acceptability as well as signal quality. Further study is required to determine the relative feasibility and accuracy in obtaining the breathing signal from these various sites. We did not find significant differences between health groups and their device form preferences.

We found that young participants were more likely to use the device for exercise, but we do not know the reasons why older people were less likely to use such technology for exercise. This could be due to overall lower exercise rates in older people or to less engagement or familiarity with exercise tracking.

Cost can be a barrier to the uptake of monitoring devices, but more than half of our participants would be happy to pay up to Aus $100 (approximately US $80) for a wearable that tracks breathing rate. At this price point, such a breathing wearable would be comparable to lower end activity trackers currently on the market and would require a simple design. While creation of a wearable is feasible at this price point, sacrifices in both reliability and comfort may arise. One area of cost reduction could be eliminating a display from the wearable. Any display could be viewed on an external screen such as a mobile phone, while alerts could be processed locally on the device.

Another consideration is device battery life, that is, power consumption must be carefully managed as a small form factor places constraints on battery life [26]. We found device charge time was negotiable, while device use time should be maintained at a minimum of 24 hours. With the size constraint of a wearable, providing this power may be difficult [6]. However, given that the majority of younger participants would like to synchronize data to their mobile phones or tablet, designers may be able to shift data processing functionality to the phone. Furthermore, since participants would like at least a week’s worth of data capacity on the device, the requirement for continuous data transmission may also be reduced.

Given the user requirement for data synchronization and data storage, it is recommended that any wearable device should primarily capture and store data. Data transmission to a mobile phone or tablet can take place secondarily by participant demand or when local device storage is full. Any advanced data processing should also take place post transmission.

User security or privacy could potentially be compromised by continuous monitoring [27,28]. We investigated privacy as an unappealing factor in this study but found no observable difference between those willing or unwilling to adopt a breathing monitor. A sample size of 10 for those who would not adopt the device prevented our analyzing a statistically significant difference between the “willingness” groups.
Limitations

There are factors limiting the applicability of our findings. The first relates to whether the sample was representative of the population in general. There was a relatively high level of technology use over the population sampled, though only a third of participants were specifically current users of health monitoring devices. Also, 60% had a university education, a high percentage of respondents were female, and the ages of the study sample were not normally distributed. Although we measured educational level, we did not measure the health literacy of the participants, which may have impacted their responses to the survey. These demographics may not be representative of the general population, and there may have been a selection bias in those who chose to complete the survey (e.g., 24% of those invited from the volunteers database agreed to participate). While we acknowledge there is a potentially high selection bias in those who chose to complete the survey towards those who were already motivated to adopt a wearable, the primary aims of the survey included determining specific user motivation and their preferences for usage and features they wish to have in such a wearable. The population captured was arguably the most appropriate to answer those questions.

Second, while we were able to show differences in the survey responses of those with and without asthma, people without asthma were younger than those with asthma, making it difficult to disentangle the effects of age and disease status. There is some suggestion that older users are more ready to adopt health-related technologies, but the reasons for this require further investigation [29]. More than half of participants with asthma also had suboptimal asthma control.

Third, display preferences were examined in a rudimentary manner in this survey, to determine whether graphical displays were preferred over text. Furthermore, we did not assess in detail whether participants understood how the information was presented, for example, by asking whether they thought the display indicated that their breathing was stable. Once wearable technology is established to measure breathing over time, another study to determine a suitable display of information from the participant’s perspective should be explored.

Finally, we did not collect data on whether those who used other health monitoring devices were current or former users, or the reasons for discontinuation of use. Information on how long and why people stay engaged beyond curiosity would have provided major insight into user psychology as well as device development.

Conclusions

We have explored the motivations for, and the likelihood of, adopting wearable technology for the purpose of breathing monitoring and identified user preferences for key design features. We found participants were motivated to adopt a wearable breathing monitor regardless of health status, yet there were distinctly different rationales for use between those with and without asthma. There is a clear need to identify the benefits of monitoring breathing in health and asthma. Next steps will require the development and testing of reliable breathing metrics or indicators that can be safely used by people with asthma for monitoring breathing over time or that assist in the identification of symptom worsening and asthma exacerbations. These findings will help inform the design of a user-acceptable wearable device that will facilitate its eventual uptake in both healthy and asthma populations.

Acknowledgments

The authors acknowledge the role of the Google Impact Challenge Grant, which provided seed funding for the greater study. We also acknowledge the role of Michele Goldman of Asthma Australia in co-development of the proposal leading to the seed funding.

Conflicts of Interest

JBP was partially funded by an Australian Postgraduate Award and a philanthropic Google Grant, of which PMY and ET were primary investigators in collaboration with Asthma Australia. In the last 3 years, the Woolcock Institute of Medical Research has received independent research funding from AstraZeneca and GlaxoSmithKline for asthma research carried out by JMF.

Multimedia Appendix 1

Online questionnaire.

References


Abbreviations

ACT: Asthma Control Test
Measurement of Skin Induration Size Using Smartphone Images and Photogrammetric Reconstruction: Pilot Study

Ronald Dendere¹, BSc (Eng), MSc, PhD; Tinashe Mutsvangwa¹, BSc (Eng), MSc, PhD; Rene Goliath², RN, BSc (Hons); Molebogeng X Rangaka²,³, MBChB, MSc, PhD; Ibrahim Abubakar³, PhD, FRCP; Tania S Douglas¹, BSc (Eng), MS, MBA, PhD

¹Medical Imaging Research Unit, Division of Biomedical Engineering, University of Cape Town, Cape Town, South Africa
²Wellcome Centre for Infectious Diseases Research in Africa, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa
³Institute for Global Health, University College London, London, United Kingdom

Abstract

Background: The tuberculin skin test (TST) is the most common method for detecting latent tuberculosis infection (LTBI). The test requires that a patient return to the health facility or be visited by a health care worker 48 to 72 hours after the intradermal placement of tuberculin so that the size of the resulting skin induration, if any, can be measured.

Objective: This study aimed to propose and evaluate an image-based method for measuring induration size from images captured using a smartphone camera.

Methods: We imaged simulated skin indurations, ranging from 4.0 to 19 mm, in 10 subjects using a handheld smartphone, and performed three-dimensional reconstruction of the induration sites using photogrammetry software. An experienced TST reader measured the size of each induration using the standard clinical method. The experienced reader and an inexperienced observer both measured the size of each induration using the software. The agreement between measurements generated by the standard clinical and image-based methods was assessed using the intraclass correlation coefficient (ICC). Inter- and intraobserver agreement for the image-based method was similarly evaluated.

Results: Results showed excellent agreement between the standard and image-based measurements performed by the experienced reader with an ICC value of .965. Inter- and intraobserver agreements were also excellent, indicating that experience in reading TSTs is not required with our proposed method.

Conclusions: We conclude that the proposed smartphone image-based method is a potential alternative to standard induration size measurement and would enable remote data collection for LTBI screening.

doi:10.2196/biomedeng.8333

KEYWORDS

tuberculosis; skin tests; telemedicine; computer assisted diagnosis

Introduction

Tuberculosis (TB) is one of the leading causes of death among infectious diseases worldwide. The disease progresses through a continuum of infection stages in individuals infected with Mycobacterium tuberculosis bacilli (MTB) from the latent to the active state. In latent tuberculosis infection (LTBI), the bacilli are largely dormant but can produce a detectable immune
reaction. The LTBI state indicates previous infection and is strongly associated with an increased risk of progression to active TB, particularly in young children [1,2].

The most widely used method to detect LTBI is the tuberculin skin test (TST), a proxy measure for previous exposure to MTB [3]. TSTs are also used to monitor and estimate prevalence of TB infection in communities [4]. During the administration of a TST, tuberculin purified protein derivative is injected intradermally in the patient’s arm, approximately 3 to 4 inches below the elbow, and the result is assessed after a 48- to 72-hour period. This means that it is necessary for the patient to return to the health facility for assessment of the outcome. During the second visit, a clinician measures, using a ballpoint pen and ruler, and records the size of, the skin induration, if there is any; the result is classified as positive or negative based on consensus thresholds. One of the problems encountered by clinicians is that some patients who are subjected to the TST do not return to their respective health facilities after the specified time to have the result evaluated. Self-assessment of the outcome of TSTs has been explored as an alternative to the required follow-up visit [5-8] and, if accurate, could enhance TB screening efforts by reducing the number of subjects whose readings are not taken because of failure to return for skin test reading [6-8]. However, patients may require training in reading TSTs [5].

As an alternative to direct clinical measurement of TST indurations, we proposed a novel smartphone-based solution, which could capture images of the induration using a camera phone and send these to a central processing center where automated or manual analyses can be performed. Mobile phone usage, particularly smartphones, and global mobile network coverage have risen sharply within the past decade, particularly in developing countries where subscription tripled from 30% to 90% between 2006 and 2014 [9]. At the same time, there was a worldwide decrease in the cost of mobile broadband between 2013 and 2016, with developing and least developed countries experiencing higher reductions than developed countries [10]. These factors increasingly make telemedicine a realistic solution. Telemedicine involves the collection of medical data in one area and transfer to a central processing station for expert analysis; it stands to benefit from the advances in imaging and networking technologies for smartphones. Smartphones are now equipped with powerful cameras and processors, large screens, and various network capabilities [11].

In this study, we addressed the measurement of skin induration resulting from a TST using images captured on a smartphone, with a vision of leveraging the ubiquity of smartphone usage and the capabilities of such phones to enhance the screening for latent TB using TSTs. We assessed the feasibility of measuring induration size for a TST in three-dimensional (3D) scenes reconstructed from spatial images acquired using a smartphone based on agreement with the standard method.

**Methods**

**Image Acquisition**

Images were captured using the primary camera of a Samsung Galaxy S7 Edge (Samsung, South Korea) smartphone to produce a set of images per subject. This smartphone has a 12-MP primary camera with a 1.4-µm pixel size, 1/2.5-inch sensor size, and a f/1.7, 26-mm lens. During image capture, particular caution was exercised to ensure sufficient overlap of the regions in adjacent images. This overlap is essential for the 3D scene reconstruction.

Subjects for the study were recruited from postgraduate students and staff members in the Division of Biomedical Engineering at the University of Cape Town. Ethics clearance for the study was obtained from the Human Research Ethics Committee of the University of Cape Town (HREC REF: 250/2016). The purpose and procedure for this study were explained to all subjects who thereafter signed a form consenting to take part in the study. Instead of subjecting each participant to a TST, we used special effects makeup, applied by a professional make-up artist, to mimic a positive TST outcome (induration). We recruited participants with different skin tones, varying from dark to pale. Each mock induration was carefully tailored to mimic the expected appearance of a real induration for that particular skin tone. Examples of the mock indurations are shown in Figure 1. For each mock induration, between 7 and 10 images were captured ensuring coverage of approximately 120° around the arm with the smartphone camera set to autofocus mode. The 120° angle was sufficient to capture the data required for full 3D reconstruction of the mock induration. Before image acquisition, we placed a 10-mm scale bar on the arm, close to the mock induration, for the calibration of the measurements. After image acquisition, the induration size was manually measured by an observer who is experienced in reading TSTs, and these measurements were regarded as the reference standard for assessing the image-based measurements. The manual measurements were recorded by an independent observer to eliminate potential bias (trying to recall the manual measurements) and ensure independence when the experienced TST reader performed the image-based measurements.

**Figure 1.** Examples of the mock skin indurations produced using special effects makeup.

http://biomedeng.jmir.org/2017/1/e3/
Three-Dimensional Reconstruction and Measurement

Images were transferred from the smartphone to a personal computer for processing using Agisoft PhotoScan (Agisoft LLC, Russia), a commercially available software package that performs photogrammetric processing of digital images. The program can generate 3D spatial data to enable indirect distance, area, and volume measurements of objects of various scales [12]. A mask was manually created for each image to ensure that only the relevant part (background cropped to leave only the arm) was used in the reconstruction. The first stage of 3D reconstruction is the search for and matching of common points on the input images; these points are used to estimate camera positions for each image. Successful completion of this stage requires points to be visible in at least 2 images, and therefore, there is a need for sufficient overlap in adjacent images. The software provides a fully automated workflow and subsequent stages include the refinement of the camera calibration parameters, building of the point cloud model, building of a polygonal mesh, and finally, building of texture. An example of this pipeline is shown in Figure 2. Identification of the scale bar in the reconstructed 3D scene was achieved by the manual placement of markers using mouse clicks. Measurements were subsequently made by placing markers on the arm to indicate the distance to be measured—on the border of the mock induration in the direction transverse to the length of the arm (Figure 3).

Evaluation

Marker placement for both the scale bar identification and the induration measurement was performed by the same observer who made the reference measurements. Additionally, an observer with no prior experience in reading TSTs also placed markers on the 3D arm models; this would allow assessment of the effect of experience on image-based measurements. The 2 observers repeated the image-based measurements so that the reliability of measurement could be assessed. Observers took the image-based measurements separately to avoid bias, and the second measurement for each observer was taken 7 days after the first one. Statistical analysis of the data was performed using the SPSS (IBM Corp, USA) software package.

Figure 2. Illustration of the three-dimensional reconstruction process.
Results

A total of 10 volunteers took part in the study. Mock induration size ranged between 4.0 and 19 mm as measured by the experienced observer according to the standard method used in clinical evaluation of TSTs. First, we studied the agreement between the readings taken by the experienced observer (observer 1) using the standard clinical method and the corresponding image-based measurements by the same observer. In this study, the precision of the clinical standard method was 1 mm (i.e., measurements were rounded to the nearest millimeter where necessary), whereas the precision of the image-based method was higher than a thousandth of a millimeter. We also assessed the agreement between the clinical standard measurements and the image-based measurements performed by an inexperienced observer (observer 2), as well as the agreement between the two observers for image-based measurements. Tables 1 and 2 show the intraclass correlation coefficients (ICC) values and their corresponding 95% CI for the various measurements conducted by observers 1 and 2.

Tables 1 and 2 show that ICCs for both observers are well above 0.9, indicating excellent agreement between the observer versus the reference and for intraobserver measurements.

Table 3 shows excellent interobserver agreement with the ICCs higher than 0.9. The ICC shown is computed using the average of the two readings made 7 days apart, by each observer.

Table 1. Intraclass correlation coefficient for measurements performed by observer 1 (experienced tuberculin skin test reader).

<table>
<thead>
<tr>
<th>Comparison</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference versus mean measurement</td>
<td>0.965 (0.865-0.991)</td>
</tr>
<tr>
<td>R1b versus R2c</td>
<td>0.989 (0.958-0.997)</td>
</tr>
</tbody>
</table>

aICC: intraclass correlation coefficient.
bR1: reading 1.
cR2: reading 2.

Table 2. Intraclass correlation coefficient for measurements performed by observer 2 (inexperienced tuberculin skin test reader).

<table>
<thead>
<tr>
<th>Comparison</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference versus mean measurement</td>
<td>0.954 (0.830-0.988)</td>
</tr>
<tr>
<td>R1b versus R2c</td>
<td>0.973 (0.897-0.993)</td>
</tr>
</tbody>
</table>

aICC: intraclass correlation coefficient.
bR1: reading 1.
cR2: reading 2.
proposed in this study does not require the patient to make any
subjects providing an interpretation. In contrast, the method
for the study were 99.5% and 97%, respectively, with all
reaction for expert evaluation [7]. The sensitivity and specificity
flat or not with an aim of summoning those with a nonflat
subjects have been asked to interpret the outcome of a TST as
TSTs has previously been tested with success. For example,
turn would also reduce the number of TSTs lost because of
practitioner, and would potentially reduce travel costs. This in
captured in their homes, without the assistance of a health
induration size indicates a negative result. Self-assessment
be required for a TST becomes optional for patients whose
size. With this approach, the follow-up visit that would normally
measured as opposed to planar distance from two-dimensional
images. The high ICC values, which indicate excellent
agreement between the measurements using the image-based
method proposed in this study and those made using the current
clinical method, suggest that the proposed method can
potentially be used as an alternative to the standard clinical
method. Moreover, high values for intra- and interobserver
agreement for both experienced and inexperienced users indicate that experience in reading TSTs is not a requirement with the proposed method.

Our method provides the ability to read the results of the skin
test at low cost, with high accuracy, in the community. An
example where this would be useful is in low-resourced, high
burden areas during mass screenings as the dependence on
experienced personnel is reduced. In such cases, personnel with
little experience or no experience in reading TSTs could be sent
to collect images for later processing. The method also
addresses lack of experience in TST reading in low-risk
countries. In high-income countries, which currently target a
wider pool of patients, the method may also provide an
alternative to clinician-based reading of results that reduces the
number of TST results that are lost because of patients’ failure
to return for the reading.

Results indicate that accurate self-assessment is possible, as
patients in possession of an adequately capable smartphone
could capture images of the skin induration and send it to a
central server for expert processing and recording of induration
size. With this approach, the follow-up visit that would normally
be required for a TST becomes optional for patients whose
induration size indicates a negative result. Self-assessment
would increase convenience for the patient as images can be
captured in their homes, without the assistance of a health
practitioner, and would potentially reduce travel costs. This in
turn would also reduce the number of TSTs lost because of
patients’ failure to return for assessment. Self-assessment for
TSTs has previously been tested with success. For example,
subjects have been asked to interpret the outcome of a TST as
flat or not with an aim of summoning those with a nonflat
reaction for expert evaluation [7]. The sensitivity and specificity
for the study were 99.5% and 97%, respectively, with all
subjects providing an interpretation. In contrast, the method
proposed in this study does not require the patient to make any
decision but to merely take images of the site where the TST
was administered. Furthermore, our proposed approach goes a
step further to produce the measurement for induration size,
and the results can be transmitted back to the patient’s phone
with the potential to include appropriate messages to encourage
attendance at a health care facility for those with a positive test.

A further advantage of the image-based method is that it
measures induration size with higher precision than the current
clinical standard method. In the standard method, a TST reader
has to round off the induration size to the nearest millimeter.
This requires the reader to make a judgment during the
measurement, and this is a possible source of variation among
TST readers that can cause misclassification (positive or
negative) of measurements that are close to the threshold [13].
The higher precision of the image-based method is likely to
reduce the misclassification rate.

Despite the relative ease with which observers could make
measurements from the reconstructed 3D scenes, a few minor
problems were encountered. First, the presence of shadows in
some of the images made the placement of markers difficult.
These shadows were a result of nonuniformity in the lighting
in the images, as the camera moves from position to position
during image capture. This can easily be addressed using image
correction techniques before 3D reconstruction. In addition,
despite close inter- and intraobserver agreement, the
measurements are still observer-dependent. One way of
overcoming this would be automated analysis, wherein an
algorithm identifies the scale bar and the points between which
the measurement is to be taken. These advances are the subject
of an ongoing study.

We identified the calibration strip, which is placed on the arm
to provide the scale, as a potential limitation for our method.
Although the placement of the calibration strip does not affect
results if it is clearly visible in the images, its physical integrity
is crucial, and therefore, it requires careful handling by the
patient. For example, patients who capture the TST images
themselves should avoid bending the calibration strip or even
losing it. Training patients on TST reading has previously been
identified as a factor for successful implementation of
self-assessment [5]; in our case, clinicians would need to educate
the patient on the importance of maintaining the integrity of the
calibration strip and demonstrate the image acquisition process
including optimal placement of the calibration strip. In this
ongoing study, we are developing a calibration technique in
line with the requirement for automated analysis, which
dispenses with the strip. Although mock indurations and a
limited number of subjects were sufficient for the current
proof-of-concept study, we envisage testing further iterations of
our method using real inductions and a larger number of

Table 3. Intraclass correlation coefficient for interobserver agreement.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>ICC² (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs1Meanᵇ versus Obs2Meanᶜ</td>
<td>0.990 (0.938-0.998)</td>
</tr>
</tbody>
</table>

ᵇICC: intraclass correlation coefficient.
ᶜObs1Mean: mean of the two readings performed by observer 1.
ᶜObs2Mean: mean of the two readings performed by observer 2.

Discussion

Principal Findings

Mock skin indurations were used in this study, and their sizes
were measured from 3D scenes of the arm reconstructed from
planar images captured at several angles using a smartphone.
The use of 3D scenery means that real 3D induration size is
measured as opposed to planar distance from two-dimensional
images. The use of 3D induration size. With this approach, the follow-up visit that would normally
be required for a TST becomes optional for patients whose
induration size indicates a negative result. Self-assessment
would increase convenience for the patient as images can be
captured in their homes, without the assistance of a health
practitioner, and would potentially reduce travel costs. This in
turn would also reduce the number of TSTs lost because of
patients’ failure to return for assessment.

Results indicate that accurate self-assessment is possible, as
patients in possession of an adequately capable smartphone
could capture images of the skin induration and send it to a
central server for expert processing and recording of induration
size. With this approach, the follow-up visit that would normally
be required for a TST becomes optional for patients whose
induration size indicates a negative result. Self-assessment
would increase convenience for the patient as images can be
captured in their homes, without the assistance of a health
practitioner, and would potentially reduce travel costs. This in
turn would also reduce the number of TSTs lost because of
patients’ failure to return for assessment. Self-assessment for
TSTs has previously been tested with success. For example,
subjects have been asked to interpret the outcome of a TST as
flat or not with an aim of summoning those with a nonflat
reaction for expert evaluation [7]. The sensitivity and specificity
for the study were 99.5% and 97%, respectively, with all
subjects providing an interpretation. In contrast, the method
proposed in this study does not require the patient to make any
decision but to merely take images of the site where the TST
was administered. Furthermore, our proposed approach goes a
step further to produce the measurement for induration size,
and the results can be transmitted back to the patient’s phone
with the potential to include appropriate messages to encourage
attendance at a health care facility for those with a positive test.

A further advantage of the image-based method is that it
measures induration size with higher precision than the current
clinical standard method. In the standard method, a TST reader
has to round off the induration size to the nearest millimeter.
This requires the reader to make a judgment during the
measurement, and this is a possible source of variation among
TST readers that can cause misclassification (positive or
negative) of measurements that are close to the threshold [13].
The higher precision of the image-based method is likely to
reduce the misclassification rate.

Despite the relative ease with which observers could make
measurements from the reconstructed 3D scenes, a few minor
problems were encountered. First, the presence of shadows in
some of the images made the placement of markers difficult.
These shadows were a result of nonuniformity in the lighting
in the images, as the camera moves from position to position
during image capture. This can easily be addressed using image
correction techniques before 3D reconstruction. In addition,
despite close inter- and intraobserver agreement, the
measurements are still observer-dependent. One way of
overcoming this would be automated analysis, wherein an
algorithm identifies the scale bar and the points between which
the measurement is to be taken. These advances are the subject
of an ongoing study.

We identified the calibration strip, which is placed on the arm
to provide the scale, as a potential limitation for our method.
Although the placement of the calibration strip does not affect
results if it is clearly visible in the images, its physical integrity
is crucial, and therefore, it requires careful handling by the
patient. For example, patients who capture the TST images
themselves should avoid bending the calibration strip or even
losing it. Training patients on TST reading has previously been
identified as a factor for successful implementation of
self-assessment [5]; in our case, clinicians would need to educate
the patient on the importance of maintaining the integrity of the
calibration strip and demonstrate the image acquisition process
including optimal placement of the calibration strip. In this
ongoing study, we are developing a calibration technique in
line with the requirement for automated analysis, which
dispenses with the strip. Although mock indurations and a
limited number of subjects were sufficient for the current
proof-of-concept study, we envisage testing further iterations of
our method using real inductions and a larger number of

Table 3. Intraclass correlation coefficient for interobserver agreement.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>ICC² (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs1Meanᵇ versus Obs2Meanᶜ</td>
<td>0.990 (0.938-0.998)</td>
</tr>
</tbody>
</table>

ᵇICC: intraclass correlation coefficient.
ᶜObs1Mean: mean of the two readings performed by observer 1.
ᶜObs2Mean: mean of the two readings performed by observer 2.
participants. Finally, in this study, we used a high-end smartphone equipped with a camera superior to those fitted on low-end phones that are expected to be more widespread in resource-limited regions. However, at this proof-of-concept stage, our goal was to show that images acquired with a smartphone can be used to measure TST induration size. Smartphone cameras have improved so much in the past few years that primary cameras on modern low-end phones have resolution of 5 MP or better. Furthermore, the 3D reconstruction software that we have used for processing the images works well with low-resolution images [12]. However, in the ongoing study, we plan to explore the effect of camera specifications on the results.

Conclusions
We have shown that a smartphone-based imaging solution has the potential to improve the efficiency of the global TB program by providing the ability to read the results of the skin test at low cost, with high accuracy, in the community. We envisage that the proposed method for induration measurement would enable and enhance latent TB infection screening in high burden, low-resource regions as well as provide an alternative to clinician-based assessment in high-income regions. By taking advantage of the global ubiquity of smartphones and through further research, the use of smartphone cameras for image acquisition has the potential to bring this method to greater portions of the population through self-assessment and application of telemedicine.

Acknowledgments
Funding for this study was provided by the National Research Foundation of South Africa (grants 82624 and 98788) and the Medical Research Council of South Africa. The funding bodies played no role in the design of the study; collection, analysis, and interpretation of data; or in writing the manuscript.

Conflicts of Interest
None declared.

References

Abbreviations

| ICC: | intraclass correlation coefficient |

http://biomedeng.jmir.org/2017/1/e3/
Heart Rate Monitoring Apps: Information for Engineers and Researchers About the New European Medical Devices Regulation 2017/745

Abstract

Background: After years in the making, on April 5, 2017, the European Parliament and Council finally adopted Regulation (EU) 2017/745, the new Medical Devices Regulation (MDR), repealing the existing Medical Device Directive (MDD) 93/42/EEC. Though long anticipated, this shift in policy will have strong and lasting effects in the medical devices industry.

Objective: This paper focuses specifically on the classification of software as a potential medical device under MDD and MDR and examines whether or not the regulatory framework for health apps has changed substantially and what, if any, impact is to expected. A particular emphasis will be on the issue of classification uncertainty raised by borderline cases such as heart rate monitoring and well-being apps. The paper primarily targets researchers and engineers unfamiliar with regulatory requirements for medical devices and aims to provide a concise, yet accurate, overview of the European regulatory framework. This is of particular relevance as with the exponential growth of fitness and health-related apps, the lines between toys, lifestyle products, and medical devices have increasingly blurred.

Methods: The recently published European Medical Device Regulation is analyzed and compared to the preceding MDD.

Results: The previous regulatory framework already provided for the possibility of apps to fall under the definition of medical devices, in which case classification rules for active medical devices applied. However, while applicability of the new regulatory framework still hinges on whether the intended purpose is medical or not, the threshold for classifying as a medical device has been considerably lowered due to a broader interpretation of what constitutes a medical purpose.

Conclusions: The adoption of the new European regulation on medical devices entails the risk that manufacturers previously unaffected by the medical devices regulatory framework may now unwillingly and unwittingly find themselves in the arena of medical device manufacturing.

Introduction

We live in interesting times. In fact, in the last 10 years or so, our already technology-driven lives became even more reliant on modern technology. Gradually and subtly we have witnessed not only our mobile phones but everyday objects become “smart.” That is, they increased in terms of functionality, connectivity, and interoperability as they evolved from common appliances to entities of what is now commonly referred to as the “Internet of Things”. Modern mobile phones harness tremendous computational power and a whole plethora of advanced sensors in the palm of a hand. Arguably, nowhere is this technological revolution more evident than in the healthcare sector where mHealth is increasingly gaining momentum [1-3].
On the other hand, it has become increasingly challenging to accurately differentiate between apps for lifestyle and well-being purposes and actual medical apps [4] (ie, apps which by virtue of their intended purpose are to be considered medical devices). While there certainly exists a risk of stifling innovation, especially in a relatively young and dynamic field such as mHealth with its many start-up companies, certainty pertaining to legal requirements and product safety classification is paramount to succeed in this highly innovative and competitive arena. Acknowledging the fast pace of innovation, it becomes increasingly important to remain up to date with the entire spectrum of legal requirements, keeping in mind that a failure to do so may likely spell the untimely demise of otherwise viable and innovative manufacturers.

This paper begins with a brief summary of the existing European Directive on medical devices, namely 93/42/EEC, its relevant amendments, and the classification of software under the Medical Device Directive (MDD). The paper then proceeds to provide an overview of the newly adopted Medical Devices Regulation (MDR) and discusses changes with respect to the MDD. The paper concludes with an assessment of the impact the new regulatory framework is likely to have on the health app sector.

The Existing European Medical Device Directive 93/42/EEC

General Background


Classification of Software as Medical Device Under the Medical Device Directive

Article 1(2)(a) MDD, as amended by 2007/47/EC, defines a medical device as:

any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, including the software intended by its manufacturer to be used specifically for diagnostic and/or therapeutic purposes and necessary for its proper application, intended by the manufacturer to be used for human beings for the purpose of:

• diagnosis, prevention, monitoring, treatment or alleviation of disease,

• diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap,

• investigation, replacement or modification of the anatomy or of a physiological process,

• control of conception,

and which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted in its function by such means.

Accordingly, the MDD already provides for software to be classified as a medical device if said software is intended to be used for one of the above listed medical purposes (eg, diagnosis, treatment). This is emphasized by Recital 5 in 2007/47/EC [6], which states:

It is necessary to clarify that software in its own right, when specifically intended by the manufacturer to be used for one or more of the medical purposes set out in the definition of a medical device, is a medical device. Software for general purposes when used in a healthcare setting is not a medical device.

Thus, given the intended medical purpose, standalone software such as (but not limited to) mobile phone apps constitutes an active medical device according to Definition 1.4 Annex IX MDD, as amended by [6], which further provides classification criteria to determine the adequate risk category within an incremental system of four classes, namely I, IIa, IIb, and III with I being the lowest and III being the highest risk class.

In particular, having established standalone software to be a medical device, Rule 10 MDD becomes fully applicable and provides that:

active devices intended for diagnosis are in Class IIa: [...]

• if they are intended to allow direct diagnosis or monitoring of vital physiological processes, unless they are specifically intended for monitoring of vital physiological parameters, where the nature of variations is such that it could result in immediate danger to the patient, for instance variations in cardiac performance, respiration, activity of CNS in which case they are in Class IIb.

Accordingly, apps for the mobile phone based detection of atrial fibrillation have already been classified as Class IIa under MDD, since atrial fibrillation is generally not an acutely life-threatening condition. If the above does not apply, apps are generally classified as Class I devices (Rule 12 MDD), assuming they are medical devices in the first place.

Note that classification as a Class I device is particularly attractive because the manufacturer can self-certify (see Annex VII MDD), avoiding the additional burden of certification through a Notified Body.

Current Situation

With the rare exception of a few Class IIa apps (eg, “FibriCheck” [7] for the detection of atrial fibrillation), fitness and healthcare-related apps on the market to date are almost entirely not labeled as medical devices due to their intended purpose, as stipulated to by the manufacturer, not being one of the medical purposes listed in Article 1(2)(a) MDD. Whether or not this will hold under MDR as well is examined in the following section.

The New Medical Devices Regulation (EU) 2017/745

Background

As the identifier 93/42/EEC implies, the MDD regulatory framework is a couple of decades old, with the respective first
drafts actually dating back to the mid-80s. The proposal for a new Medical Devices Regulation was first published in September 2012 [8], setting in motion a lengthy negotiation, review, and amendment process. The MDR was eventually adopted on April 5, 2017, published in the Official Journal of the European Union on May 5, 2017 [9], and entered into force on the 20th day following its publication (ie, May 26, 2017). However, there is a transitional period of approximately 3 years (Article 120 MDR).

**Regulation Versus Directive**

Note that as opposed to the MDD, the new MDR is a regulation rather than a directive. The two differ in that a regulation is directly enforceable in all member states, whereas a directive first requires an implementation into national law. The German Medical Devices Act (Medizinproduktegesetz, MPG) can be seen as Germany’s implementation of the MDD.

**Broader Definitions**

While similar at first glance, the new MDR in fact substantially differs from the MDD it repealed. The following appraisal will be limited to selected changes relevant to the discussion at hand since an analysis of further aspects would be beyond the scope of this paper.

A first (and arguably easily overlooked) important difference arises in the very definition of “medical device”, defined in Article 2(1) MDR as:

> any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:
> • diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease

Note the expansion of the definition of medical device (further expansions of the definition by Article 2(1) MDR will be disregarded here) to devices engaging in the prediction and/or prognosis of diseases. Accordingly, an app that collects and possibly aggregates data from various sensors and the subject’s input (think of big data and analytics) to assess a subject’s risk of developing a specific disease or condition or the likelihood of worsening or improvement of an existing disease or condition may have to be classified as a medical device under Article 2(1) MDR. However, this would require said device to be explicitly intended for the prediction and/or prognosis of disease. As discussed in more detail below, the MDR provides for an exemption for devices whose intended use lies in nonclinical arenas such as lifestyle or general well-being. However, a manufacturer is obliged to accurately and unambiguously indicate the device’s intended use without acting arbitrarily for the sole purpose of circumventing the (perhaps unfavorable) medical device classification.

**Classification of Software as Medical Device Under the Medical Devices Regulation**

Similar to MDD, MDR provides an extensive list of classification criteria (Annex VIII MDR). The significant change lies in the adoption of a specific classification rule for software not encountered previously in either the original MDD or its amendments. Rule 11, Annex XIII MDR states:

*Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa, except if such decisions have an impact that may cause:*

• death or an irreversible deterioration of a person’s state of health, in which case it is in class III; or
• a serious deterioration of a person’s state of health or a surgical intervention, in which case it is classified as class IIb.

Software intended to monitor physiological processes is classified as class IIa, except if it is intended for monitoring of vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as class IIb.

All other software is classified as class I.

While the classification criteria of MDD and MDR as previously discussed are virtually identical, the broadening of the definition of “medical purpose” in Article 2(1) MDR in combination with Rule 11 MDR has far-reaching implications. While under MDD (among others) an emphasis on diagnosis or monitoring of treatment of disease is required to fall into the scope of the Directive’s definition of medical device, under the new MDR (as previously discussed) the scope has been broadened by the addition of “prediction and prognosis” of disease. This means that many apps, which did not constitute medical devices under MDD, now become medical devices, most likely Class IIa.

**Conclusions**

With the adoption of the new European regulation, health apps are now more likely to fall under the definition of medical device and thus become subject to the provisions of MDR.

With the inclusion of “prediction and prognosis of disease” in Article 2(1) MDR, a heart rate monitoring app for instance may now be subject to MDR as soon as the heart rate monitoring functionality itself is enhanced by health assessments (ie, inferring the subject’s cardiovascular health on the basis of heart rate or heart rate variability measurements) and the like, which would have to be interpreted as a prediction and/or prognosis of disease.

However, the applicability of the new regulatory framework still hinges on whether the intended purpose of the app, as stipulated to by the manufacturer, is medical or not. Apps solely intended for lifestyle or well-being purposes represent an important exception, for they do not constitute medical devices (see Recital 19 MDR). However, it has arguably become much easier to cross from being a lifestyle product to being a Class IIa medical device, a fact that engineers and researchers must recognize.
Acknowledgments
ML was supported by the “Excellence Initiative” of the German Federal and State Governments and the Graduate School of Excellence Computational Engineering at Technische Universität Darmstadt. The views expressed in this paper are solely those of the author in his private capacity and do not necessarily reflect the views of Technische Universität Darmstadt or any other organization.

Conflicts of Interest
None declared.

References

Abbreviations
CNS: central nervous system
EU: European Union
MDD: Medical Device Directive 93/42/EEC
MDR: Medical Devices Regulation (EU) 2017/745
mHealth: mobile health
MPG: Medizinproduktegesetz (German Medical Devices Act)