Experience: Design, Development and Evaluation of a Wearable Device for mHealth Applications

George Boateng,
ETH Zurich

Vivian Genaro Motti,
George Mason University

Varun Mishra,
Dartmouth College

John A. Batsis,
Dartmouth-Hitchcock Medical Center

Josiah Hester,
Northwestern University

David Kotz
Dartmouth College

Abstract

Wrist-worn devices hold great potential as a platform for mobile health (mHealth) applications because they comprise a familiar, convenient form factor and can embed sensors in proximity to the human body. Despite this potential, however, they are severely limited in battery life, storage, band-width, computing power, and screen size. In this paper, we describe the experience of the research and development team designing, implementing and evaluating *Amulet* - an open-hardware, open-software wrist-worn computing device - and its experience using Amulet to deploy mHealth apps in the field. In the past five years the team conducted 11 studies in the lab and in the field, involving 204 participants and collecting over 77,780 hours of sensor data. We describe the technical issues the team encountered and the lessons they learned, and conclude with a set of recommendations. We anticipate the experience described herein will be useful for the development of other research-oriented computing platforms. It should also be useful for researchers interested in developing and deploying mHealth applications, whether with the Amulet system or with other wearable platforms.

Keywords

Mobile Health; Wearables; Sensing

1   INTRODUCTION

Mobile health technologies have great potential to enhance quality of life, improve public health [1, 15], and reduce health care costs [20, 28]. mHealth devices and applications are proliferating [11]. Making the challenges surrounding usability, manageability, interoperability, availability, security, and privacy more urgent [17, 27, 29, 30, 32]. In this paper, we present the experience of the *Amulet* project, which sought to lay the scientific foundations for secure, wearable, mHealth technologies. The result, the Amulet wrist-worn mHealth platform shown in Figure 1, provides an open framework for body-area pervasive computing, centered around health-monitoring and health-management applications like stress measurement, activity monitoring, and technology-guided behavior change.

A computational bracelet has potential to form the hub of a body-area mHealth network, coordinating the activities of other on-body or near-body health devices. Such devices complement the sensing and interaction capabilities of a smartphone, because they are designed to be worn close to the skin and in location(s) where health-relevant activity can best be sensed. Furthermore, a wrist-worn platform provides a discreet means for communicating with the user, more convenient for a quick glance and more likely to be present than the smartphone itself [19, 26].

The Amulet team developed the platform - including hardware, software, and toolchain - to be a flexible method supporting researchers in conducting health-related field studies. The team had three major goals: (1) support the development of energy-efficient applications, (2) secure the platform so as to preserve participants' privacy, and (3) design interfaces that are intuitive for end users and practitioners.

Health researchers can develop or adapt one or more apps for the Amulet [4], using its development tools to forecast the battery life and memory consumption of their apps. These researchers can then field the Amulet and apps to volunteer participants, for days or weeks, with data accumulating on the internal micro-SD card. Indeed, several of our research teams have deployed the Amulet in studies, primarily on two topics: stress monitoring and physical-activity monitoring. Most of the studies collected data via a combination of Amulet's in-built sensors, other wearable sensors, and ecological momentary assessment (EMA) questions posed on the Amulet screen [7, 12, 22].

In this paper, we focus on the five-year experience of the Amulet project, combining perspectives from multiple stakeholders: developers, designers, study participants, health practitioners, health researchers, student and staff researchers, and the principal investigators. Sixteen investigators have been intimately involved during the project: 9 computer scientists, 2 electrical and computer engineers, 2 healthcare practitioners, and 3 experts in human-centered computing. We collected data on the stakeholders' experience by asking them to reflect on their work and by surveying them online. Our analysis of the stakeholders' experiences includes the challenges they faced as well as the approaches they either employed, or recommended, to address potential obstacles.

Despite extensive research on wearable health platforms, focusing on specific devices, patients, and medical conditions, we are unaware of any published work that explores the *experience* of these research groups in a way that is accessible to, and useful to, other mHealth or wearable-computing researchers. In particular, recent "experience papers" at MobiCom have focused on very different topics - such as smartphone roaming behavior in European countries [21], the performance of commercial devices concerning respiratory monitoring [14], and using Android phones as gateways in Zanzibar [16].

After an introduction to the Amulet platform, and the hardware platform evolution, we describe eleven studies that have been conducted with the Amulet. We describe some of the challenges the team encountered when developing and deploying the Amulet, and summarize a set of lessons learned.

2   AMULET PLATFORM

The Amulet project envisioned *computational jewelry,* in the form of a bracelet, that could provide the foundation for mHealth research and interventions [34]. The resulting platform comprises the hardware and software for the bracelet, as well as a suite of development tools. The entire platform, including hardware and software, is open-source and available for download on GitHub. The Amulet may be freely reproduced for research and education purposes.

The Amulet design evolved iteratively from 2012 to 2018 using off-the-shelf components on a custom Printed Circuit Board (PCB). The evolution of the Amulet is shown in Figure 2, with the major changes listed. The current Amulet device is built around two microcontrollers: an MSP430 running applications, and an nRF51822 for communicating with Bluetooth Low Energy (BLE) devices such as a heart-rate monitor, an exercise device, or an electro-dermal activity (EDA) sensor. Built-in sensors measure acceleration, rotation, ambient sound, ambient light, and ambient temperature. End users can interact with the applications via two buttons and a capacitive touch slider. A haptic buzzer, two LEDs, and a low-power high-contrast LCD display are used for visual and haptic feedback. A micro-SD card reader provides up to 2GB of internal storage. With an 110 mAh battery the Amulet device can last days or weeks, depending on the apps.

The Amulet device is coupled with a software framework that enables developers to create safe, secure, and efficient mHealth applications that fit seamlessly into everyday life [13, 34] - together they comprise the Amulet Platform. The software framework includes the Amulet Firmware Tool- chain - which builds application firmware - the Amulet Runtime system, and the Amulet Resource Profile viewer (ARP-View) graphical tool, which allows developers to interactively explore the impact of implementation decisions on battery life.

2.1   Why was Amulet Needed?

There are several wrist-worn commercial and research devices that are used for mHealth purposes; this was true even at the start of the Amulet Project [18]. These devices can be categorized into two types: single-application and multiple-application devices. Single-application devices tend to have a specific health goal such as step counting and sleep

tracking. Examples include Fitbit Fitness Trackers (such as: Flex and Charge), Jawbone, and Garmin. These devices tend to have longer battery life, running highly-tuned proprietary code. Multi-application devices have the capability to install and run several applications; examples include smart watches like Fitbit Versa and Ionic, Pebble, Apple Watch, and Wear OS (formerly named Android Wear). These devices include some pre-installed apps, and allow development and installation of new custom apps. These devices tend to have a short battery lifetime of a few hours or days.

Although single-application devices may have long battery life, they run proprietary algorithms and researchers cannot adjust or extend the set of apps on the device, limiting their usefulness for research. Also, researchers cannot modify the applications to a target population since these commercial devices are meant to evaluate populations as a whole, rather than individual specific groups. The Amulet fills this gap, providing the advantages of both single-app devices (long battery life) and multi-app devices (flexibility) while allowing researchers to explore underlying systems issues as well (because it is open source, open hardware).

Long battery life is especially important for populations like the older adults we have targeted for several of our mHealth studies [3]. Older adults tend to cite inconveniences with using health technologies that require physical and cognitive efforts [24]. Indeed, in our own mobile-health studies we found that device abandonment increases as device management burden (such as charging devices) increases. Devices that require infrequent charging can reduce the burden of participation [7] and encourage ongoing participation. In some populations, zero-effort device management is critical (for example, pediatric populations or geriatric dementia populations).

2.2 Amulet Design Requirements

In spite of the severely-constrained power resources of wearable jewelry, the original Amulet design aimed to address flexibility, efficiency, reliability, security, privacy, and usability, among other quality requirements. The Amulet vision was to enable longitudinal experiments on human subjects in a wide variety of health-behavior science research, and to demonstrate that it is feasible to deploy a secure, multiapplication wrist-worn platform that can run an mHealth workload for days or weeks on a single battery charge. In pursuing this vision, the team developed a series of iterative prototypes, folding experience from early pilot studies into improved hardware and software design.

A multidisciplinary team of investigators collaboratively developed the requirements for the 11 user studies described in Section 3. They based their decisions on the study purpose, number of participants, logistics, duration, location, data collection, and analysis. The research team included expertise in Engineering, Computer Science, Human Factors, and Health Sciences. The team found three criteria for success, common across several deployment studies, defined as follows.

Non-intrusive: The system and application designs should minimize intrusion, so that study participants could still perform their everyday activities without being disturbed by the application. Furthermore, the bracelet should be lightweight, fit comfortably, provide

convenient access to review alerts from the app, and allow easy readability and fast input by pressing or holding buttons.

Long battery life: The system and applications should be designed for long battery life, to reduce the risk of incomplete data collection or data loss due to device failure. The team chose software and hardware designs for reliability and efficiency, and implemented a stand-by state to save power when the user is not interacting with the device.

Flexible and rich sensing: The device should support a diverse range of applications, by incorporating a rich set of sensors inside the device and supporting access to other sensors via BLE. Apps should have access to raw sensor data when desired. In contrast with most commercial devices, which limit app developers to low sensing rates or to high-level inferences, Amulet allows researchers to innovate by creating new algorithms for sensor-data processing.

2.3   Amulet Hardware Revisions

As shown in Figure 2, the Amulet device went through significant changes as technology improvements, manufacturing ability, and the evolving needs of the project demanded. Often these changes were in response to lessons learned from deployment and focus groups. We detail the changes below.

Development Board: The first Amulet (Figure 2A) was built with an ARM Cortex-M4, an ANT radio (nRF51422), numerous sensors, and a small MSP430 intended for running low power background tasks. Powered by a wall wart, this original version was not wearable. The team designed sophisticated battery charging, power conditioning, and energy management circuitry into this version in anticipation of the wearable form factor. The prototype had nearly all its pins broken out in a large format for simplified testing of software functionality, and debugging of hardware issues. This board was sufficient to develop the first version of the multi-tenant runtime system (which later evolved [12]), and benchmark different configurations, as well as initial driver development and simplistic applications.

First Wearable: After initial software validation and testing on the large development board, the team reduced the design into a 1.4×1.4 inch wearable format (Figure 2B), with minor changes in pin-out and the removal of all debug headers. This design included an OLED display and a LiPo battery, enclosed in a custom 3D-printed enclosure and attached to the wrist with a clasp strap.

Two-board Design (MSP430 conversion): Development of the Amulet framework was slowed by the complexity of the Cortex-M4, licensing issues related to the proprietary IDE for embedded ARM development, lack of isolation features on the M4 (at the time), and lack of ARM expertise on the development team. Additionally, a silicon bug meant that sleep-mode energy consumption was 10x higher than reported, and the OLED display drew significant amounts of current, shortening battery lifetime to a few hours. At this point the development team also noted that the M4 might be over-provisioned for the proposed

applications. This combination of factors led to a new revision (Figure 2C), and the adoption of a 16-bit MSP430 as the main microcontroller. It had a quarter of the code space and ran at a tenth the speed; but it was cheaper, simpler, and good enough to run many applications. The team swapped the OLED display for a SHARP LCD display to reduce power consumption, and replaced the ANT radio with a BLE module. To simplify independent development of the BLE drivers (since the Amulet device was intended to function as a "Central" device that could manage other peripherals), the design spread across two PCBs connected by a cable; one PCB contained the Application processor (MSP430), the other the BLE module and some sensors. The original intent was that the BLE module, equipped with a ARM M0 that ran the BLE stack, could also manage some of the sensor data collection and storage. The design included a rotary encoder (for wheel-like touch input) to enable scrolling alongside the buttons. Finally, the team redesigned the 3D case into a watch form factor.

Fully Integrated Design: Despite the team's best efforts, the two-board design had two key downsides. First, the stacked form factor and more open mechanical design made the case very large, bulky, and unattractive according to feedback in focus groups. Second, the separation of program concerns across the two processors, made for a distributed debugging headache. To gather data from the IMU, the Application processor had to initiate the inter-processor-communication, then the BLE processor had to collect data and send it back. This was complicated by the interrupt driven and timing sensitive BLE stack (a softdevice), which generated random non-maskable interrupts that would corrupt or end data collection. This caused any non-BLE program on the processor to be unreliable. We redesigned the Amulet device from scratch focusing on simplicity, centralization of control, and size (Figure 2D). The main processor hosted all application and operating system functions, and treated the BLE module as a modem. The team simplified and miniaturized the power and energy management circuitry, partially due to technology advancement, partially from removing superfluous modules like a coulomb counter. A high-contrast version of the LCD replaced the older version, making text easier to see. Additionally font sizes were increased and a new display module API was developed to facilitate a richer user interface. Instead of the rotary encoder, the team used a five-input capacitive touch pad. This change reduced the footprint of the device and made for a more comfortable scrolling interaction. A new mechanical design made with flexible silicone provided a lower profile, and fit common 20mm watchbands. This version, the "White Owl", was the first version successfully deployed in multiple studies, and was presented at SenSys 2016 [13].

Current Version: The team redesigned the Amulet for comfort, looks, and technology evolution. The new case has rounded corners and a miniaturized design. The team added a UVA/B sensor to enable new applications studying sun exposure. They replaced the USB plug with a through-hole version that increased mechanical reliability. They upgraded the MEMS-based sensors for power savings, and replaced the Bluetooth module with a chip that supports BLE5.

Price Evolution: Over the years each Amulet device cost varied from $500 to $200, including all components, assembly, enclosure, battery, and screen. The most recent Amulet,

thanks to technology improvements and reduction of component count, costs U$169.26 per unit in a batch of 100, manufactured in the United States.

3 AMULET-BASED USER STUDIES

The Amulet team and its collaborators conducted 11 studies (excluding focus groups) in the past five years: some in controlled environments (laboratory settings) and some in free-living conditions (in the field). The studies, hardware revisions, and source releases are shown in the timeline in Figure 3. In all, 204 young and older adults participated in these studies, as shown in Table 1. These studies involved a variety of sensors, including Amulet's built-in sensors (accelerometer, gyroscope) and wearable external sensors (heart R-R interval (RRI) and electro-dermal activity (EDA), also known as galvanic skin response (GSR)), as well as ecological momentary assessment (EMA) questions posed to the participant on the Amulet screen. Data collected in the studies was qualitative and quantitative, resulting in over 77,780 hours of data in total. All studies were approved by the Institutional Review Board of the participating institutions.

We summarize the experience of the Amulet team and its collaborators - and, indirectly, the participant experience - in the following sections.

3.1 Amulet: Human Factors and Design

We conducted studies with current and prospective users prior to (and during) the design of the hardware and form factor. To understand the preferences of current users about wearables, we conducted interviews and analyzed online reviews (available in e-commerce websites) about commercial devices, including fitness trackers and smartwatches. These studies provided a list of key wearability requirements for the Amulet [25], helping to understand the problems users faced during interaction, and their preferences in terms of functionality [19]. The requirements identified concerned device unobtrusiveness, wearer comfort and privacy. Although these user studies were based on existing off-the-shelf trackers (such as Jawbone, Fitbit, Polar) [19], they nonetheless provided prospective solutions eliciting the preferences of end users for next-generation wearables. Those studies were valuable to clarify users' preferences for interaction modes, battery life, and interface design, particularly for specific user populations (such as older adults, athletes and students with disabilities). They also resulted in recommendations for user interface design [26] and privacy-preserving controls [27]. Such contributions laid the scientific foundations for further work on the domain of wrist-worn interaction. Still, we note that it was not feasible to address all wearability requirements in the design of the Amulet prototype, mainly due to resource constraints (personnel, time and financial costs) and inherent trade-offs between wearability and functionality.

We also conducted three focus groups early on with prospective wearable users (n=24 participants) and 30 interviews with users who already had experience with wearables. Our key findings clarified what users wanted in terms of interaction, modalities for input entry and output responses, design of graphic user interfaces, battery life and recharges, etc. We briefly describe: (1) what user studies were conducted, (2) the key findings about wearability, and (3) how the team addressed users' recommendations in the Amulet.

Findings from Focus Groups: Generally, the 24 participants preferred wrist-worn devices, with armbands and smart watches being popular. The preferred functions included fitness monitoring, healthcare and communication. Input and output modalities varied, but tactile interaction was the most frequently requested, combining (for instance) vibration and touch. For all devices, a combination of modalities was preferred. The key quality factors involved size (small, minimalist, thin), flexibility (fashionable, customizable, and adjustable) and style (conventional, sporty, stylish).

Findings from Interviews: Current users, though generally satisfied with their devices, reported a number of problems with their interaction and provided suggestions for improvements. Preferred form factors were small, light weight, comfortable, and easy to use. A combination of modalities is preferred for the user interaction, especially tactile (e.g., touchscreen and vibration) and auditory modalities. Participants' motivations to use wearables vary, but most related to healthcare and fitness or for communication.

Most interviewees preferred a wrist-mounted device. Their preferred interaction varied depending on personal preference, the device purpose, or the circumstances of the interaction (context). Overall, audio interaction was seen as beneficial due to its hands-free property, but if not carefully designed, audio can also be annoying, disturbing, and not private for wearers. Vibration was considered as positive to reinforce information, for instance when users hit a goal. Vibro-tactile interaction was considered to be more private and discreet than audio, and suitable for wrist-worn wearables. Users preferred touch-screen, tactile displays and physical buttons, but the interaction needs to be carefully designed to provide interfaces that are effective, minimalist and simple to use as well.

To address the above wearability recommendations in creating the Amulet design, the research team strived for (a) a small form factor by employing small electronic components and reducing board size to the extent possible, (b) optimized battery usage, with stand-by modes, activation cycles, and efficient data processing, and (c) multiple modalities for input entry and output responses.

3.2 Studies in 2016

The Amulet team conducted two studies in 2016 for monitoring stress. According to The American Medical Association, stress is the underlying cause of more than 60 percent of all human illness and disease. Because stress is often contextual and in-the-moment, longterm in-situ monitoring and intervention seem most promising for treatment [35]. One study was conducted in a laboratory setting and one in the field [8, 9]. Both studies used the Amulet version shown in Figure 2D, nicknamed "White Owl" by the team.

Stress 1: In the first pilot study, the team collected data from two subjects in the lab [8, 9]. For about 80 minutes the researchers subjected participants to mild "stressors" (activities that previous experiments have shown to induce stress). Each subject in the protocol was exposed to six rest periods and five stressors: public speaking, mental arithmetic (twice), startle with a clap sound, and cold pressor (hand immersed in ice water). During the lab study, each subject wore a Zephyr chest strap, a commercial device that measures heart rate and RRI. RRI is commonly correlated to stress [35]. The Zephyr continuously transmitted

heart-rate and RRI data to the Amulet throughout the duration of the study. Over three hours of data were collected. Researchers asked the subjects to rate their stress level periodically (low, medium and high), which the team later used as ground truth to develop an algorithm for stress detection.

Stress 2: In this field study, the researcher collected data from ten subjects [8, 9]. The participants wore the Amulet and Zephyr for one day, during waking hours for about 8 to 12 hours. The Zephyr transmitted heart rate and RRI data to the Amulet throughout the day. The Amulet also recorded acceleration data. A custom "StressAware" app on the Amulet logged data with a duration of five minutes for some devices and one minute for others, every 10 minutes. StressAware then prompted the subjects to answer two EMA questions, asking them to rate their stress level and their activity level as low, medium or high at the moment. There were four EMA prompts per hour and at least 32 EMA prompts per day, providing data later used as "ground truth" for training and testing algorithms. This study collected over 100 hours of data, which the researchers used to develop an algorithm for stress detection.

The researchers collected usability feedback from subjects in the field study via a survey. Eighty percent (80%) of the participants in the field study agreed that it was easy to answer the questions on the Amulet, which demonstrated the potential of using wrist-worn devices such as the Amulet for EMA studies. Some expressed concerns about the bulkiness of the Amulet, and the frequent disconnections from the Zephyr. This feedback informed the redesign of the Amulet to make it less bulky. Also, subsequent studies used devices such as the Polar H7 that had a more stable Bluetooth connection with the Amulet.

3.3   Studies in 2017

The Amulet team and its collaborators conducted four studies in 2017, some focused on stress and some focused on physical activity of older and young adults. These studies used the Amulet version shown in Figure 2E, nicknamed "Kite" by the team.

Stress 3: Researchers collected data from 27 participants in a study to measure stress [23, 33]. Given the concerns with the Zephyr in the previous stress studies, the researchers took a conservative approach towards selection of the heart-rate monitor. They compared the performance of the two most-popular heart-rate monitors available at the time: Zephyr HXM and Polar H7, along with the Amulet, for reliability and validity of data. They conducted a preliminary test to compare these heart-rate monitors with a popular clinical ECG device, the Biopac MP150 [5]. They first measured participants with the Zephyr and the Biopac, and then with the Polar H7 and the Biopac. Correlation analysis of the popular features computed from the data collected, revealed that the features from Polar H7 had a stronger correlation with the Biopac (on average, $r > 0.9$, $p < 0.001$), as compared to the correlation between the features computed from the Zephyr HXM and Biopac (on average, $r = 0.6$, $p < 0.001$). Given the better performance of the Polar H7, they used it for the study [23].

Study participants wore the Amulet on one wrist, the Polar H7 on their chest, and a custom EDA sensor on the other wrist. They spent an initial 45 minutes in a controlled lab experiment, where they were exposed to three stress-inducing tasks. Following the lab

session, the participants wore the devices in the field for three days for at least 8 hours per day. In the field, the Amulet posed Ecological Momentary Assessment (EMA) questions about the participants' stress levels and mood. The study collected heart rate, R-R interval, EDA, accelerometer, EMA, and survey responses, accumulating over 1 GB of data. The study collected 668 hours of sensor data, and 1,246 valid EMA responses. The team used this data to develop algorithms for stress detection. Each EMA question prompted for a response on 5-level Likert scale, which were later scaled down to a binary outcome, to be consistent with the lab sessions. The researchers also conducted a qualitative and quantitative usability survey and observed that the population of this study (college students between 18–28 years) found the Amulet easy to use, and could have worn the Amulet and the other sensors for a longer period of time. While participants did not find the Amulet to cause physical discomfort, some felt self-conscious wearing the Amulet in public.

Activity 1: In a semi-controlled study, the researchers collected accelerometer data from 29 subjects as they wore the Amulet and performed various physical activities: lie down, stand, sit, walk, walk fast, and run [6, 10]. Researchers recruited two cohorts: young adults (n=14) and older adults (n=15). The young adults were college students ranging from 18 to 23 years old and the older adults were all above 65 years old. Researchers collected data for the young adults at Dartmouth's gymnasium, and for the older adults at a local aging center affiliated with the Dartmouth-Hitchcock Medical Center. This study collected over 29 hours of data. There was an app running on the Amulet with which subjects specified which of the activities (lie down, stand, sit, walk, walk fast, and run) they were about to perform, indicating the start and end of the activities they performed. The researchers later used this information (about the specific activities corresponding to the collected data) as ground truth. The Amulet collected acceleration data from its embedded Analog Devices ADXL362, a commercial and validated accelerometer. The researchers used the data and the ground-truth information to develop an activity-detection algorithm. The results of this study allowed the investigators to later test this algorithm outside of a controlled lab setting (see Activity 2 and Activity 3 studies below).

Activity 2: Researchers ran a five-day field study in which five older adults (ages: 73, 73, 83, 86 and 87 years) each wore an Amulet as it monitored his or her activity level [7]. This "GeriActive" app tracked how much time they spent doing low, moderate, or vigorous activity, and the duration the Amulet spent in a non-wear state. The app also tracked battery life and logged an hourly summary for later analysis. The app collected a total of 120 hours of data. The app provided subjects a display of their progress toward a daily activity goal, and delivered encouragement alerts three times a day. Usability data from subjects showed a general satisfaction with the Amulet and the GeriActive app.

Activity 3: In this study, seven older adults wore an Amulet for three days in the field as it ran the "GeriActive" app [7], and tracked the number of minutes they spent doing low, moderate and vigorous activity. The app asked subjects to respond to EMA prompts periodically, in which they were asked to indicate their activity level in the recent minute. The app collected data from the accelerometer, and battery level, with over 504 hours of data collected. The goal was to validate the GeriActive algorithm in the field. Subjects also

completed an exit questionnaire. This study permitted cross-validation of EMA and activity with documented objective activity in an older adult population. The study made clear some of the challenges in designing for usability with older adults; including screen readability, need for tactile feedback and higher contrast displays, and the necessity of reducing device management burden.

3.4 Studies in 2018

The Amulet team and its collaborators conducted four studies in 2018, focused on the physical activity of young and older adults. Studies in this year also used the "Kite" shown in Figure 2E.

Steps 1: Researchers collected accelerometer data from subjects as they performed different kinds of walking activities. Participants included younger adults (n=20) and older adults (n=20, age above 65 years). This study collected over 20 hours of accelerometer data along with EMA responses. The team used the data to develop and validate a step algorithm and a pedometer app for the Amulet.

Steps 2: In this study, twelve subjects consisting of younger adults (n=5) and older adults (n=7, age above 65 years) wore an Amulet and a Fitbit on the same wrist for two days. The study collected an hourly record of steps taken resulting in 576 hours of data. Researchers ran this field study to validate a previous Amulet-based step algorithm and pedometer app against a Fitbit device.

Strength: To assess strength training, the Amulet team built a custom force-measurement handle for commercial "Thera-band" resistance exercise bands [31]. The researchers ran a one-day lab and field strength-training study with 12 younger and older adults in one cohort, 20 older adults in a second cohort, and 8 older adults with obesity in a third cohort. The subjects performed four strength-training exercises with the customized resistance band, which sent data to an app on the Amulet [2, 31]. The app tracked the force exerted throughout the exercises and displayed summarized information on the Amulet screen. This study collected over 16 hours of data. The investigators had positive experiences using Amulet as a test platform for application development and capturing of such data. This experience was useful to researchers and clinicians alike, permitting an iterative, user-centered design process.

Weight Loss 1: Researchers ran a four-month field study in which 16 younger and middle-age adults each wore an Amulet as it ran an app, "ActivityAware" [6, 10], which monitored their level of physical activity. The ActivityAware app tracked the number of steps of users as well as the amount of time they spent doing low, moderate, or vigorous activity, and also the amount of time the Amulet spent in a non-wear state. The app tracked battery life and logged a summary of this information hourly for later analysis. The app presented subjects with a display of their progress toward achieving their daily activity goal, and delivered encouragement alerts three times a day. This study collected over 43,000 hours of data. Unlike the earlier field-based studies, qualitative data from this cohort demonstrated

that patient participants involved in a clinical care setting wished to have a more aesthetically pleasing device with more accurate measurements.

Weight Loss 2: Researchers ran a three-month field study involving 16 older obese adults (above 65 years). Each study participant wore an Amulet running the "GeriActive" app [7] to monitor their level of physical activity. The GeriActive app tracked their number of steps and the amount of time they spent doing low, moderate, or vigorous activity, in addition to the amount of time the Amulet spent in a non-wear state. The app tracked battery life and logged a summary of this information hourly for later analysis. The app presented subjects with a display of their progress toward achieving their daily activity goal, and delivered encouragement alerts three times a day. Over 32,000 hours of data was collected during the course of this study by the Amulet devices. Generally, older adults were much more willing to wear the device for long periods in the field compared to younger adults.

4 CHALLENGES

We sought to understand the entire Amulet team's experience and diverse perspectives, thus we conducted an online survey. We collected information about the team member's role and duration of his/her participation in the project, and asked them to summarize the challenges they faced, lessons learned, and recommendations. Nine stakeholders (n=9) responded to the survey, including the authors of this experience paper. Team member responses came mostly from graduate students who worked as research assistants (n=5), followed by investigators (n=3), and senior research programmers (n=1). All participants had worked in the project for longer than one year. We summarize their responses below.

Most respondents (n=7, 77.8%) were involved in software specification and development, followed by hardware test and evaluation (n=6, 66.7%), software test and evaluation (n=5, 55.6%), and planning and conducting user studies (n=5, 55.6%). Additional activities included data collection and analysis (n=4, 44.4%), team management (n=4, 44.4%) and hardware design and implementation (n=3, 33.3%).

When questioned about the major challenges they faced, respondents commented on the broad specifications (hardware and software decisions that serve a broad range of applications), dynamic updates of the code, limited documentation, difficult debugging, lack of backward-compatible hardware, lack of debug logs, lack of prior benchmark studies with baseline estimates on battery life, limited processing capabilities, and limited user-friendliness for end users as well as programmers. We categorized identified challenges in four high-level themes: system, usability, integration and deployment, described as follows.

4.1 System

Stakeholders commented on various implementation aspects of the system, notably the software (coding, debugging, and documentation), as well as processing power and memory.

Code, debugging and documentation: Respondents commented on the limited documentation for the built-from-scratch platform, which led to a steep learning curve for new team members to learn how to use the framework. Because the entire implementation

occurred in a university setting, key personnel (PhD students) would graduate and not always document best practices or new features implemented in the pursuit of publication. Respondents noted challenges in debugging applications because of the multi-chip device and the intricate combined ARM, MSP430, and JAVA toolchain that the platform was built on. The Amulet OS did not provide logs when there were bugs in the OS, which made it difficult to track the cause of strange bugs - especially when Amulet devices returned from the field. Another respondent noted that errors were not always reproducible, making it difficult to identify the causes of inconsistent behavior, or from which layer of the system stack the error originated. Respondents suggested a better reinforcement of best practices regarding coding standards and tools. They also recommended systematic development protocols, to ensure efficient management of the software.

Data Loss: Many respondents found that software issues (such as a mediocre SD card library) combined with poor hardware choice (using a lower quality and more affordable brand of SD cards) often led to data loss which could impact the validity and the results of some studies. Also, when the Amulet did run out of battery it would fail to complete data collection in some studies. Recharging the device would lead to its reset, which eventually also led to inaccurate timestamps and data loss. Team members dealt with these issues in a variety of ways including in-study debugging, reliance on alternative data to fill in gaps (including self-report), and hardware redesign.

Processing power and memory: The small size of the device and choice of platform (16-bit MSP430) limited its processing power and available code size as applications became more sophisticated than envisioned at the start of the project. This sometimes led to study-specific decisions such as running only one task at a time, and creating one application per study instead of multiple applications that could execute in parallel.

4.2   Usability

Concerning usability, Amulet team respondents reported some participants feeling uncomfortable or even ashamed to wear some versions of the prototype, and practitioners sometimes found it difficult to create new applications.

Wearability: Some participants found it difficult to accept the device or to use it comfortably, as they felt embarrassed wearing the prototype (some carried it inside a pocket or under sleeves). Building a device that serves multiple purposes proved to be difficult when a minimalist design is envisaged. In fact, the lack of experts on mechanical design, fashion and product design or aesthetics might have hindered user acceptance in general. Also, study participants in the various research studies tended to compare the prototype with commercial watches, which have a leaner design and are more aesthetically pleasant, but do not satisfy the requirements of many mobile health projects.

Ease for healthcare researchers: Stakeholders included healthcare and clinical practitioners who tried to create customized applications; they reported that Amulet documentation was complex for non-technical users to understand and use. The clinical

researchers also reported operational challenges when interacting with the device, onboarding new participants in studies, and ensuring proper device usage in the field.

4.3 Integration

Survey respondents reported challenges related to design decisions that span hardware, software, system, application, and usage.

Cross-cutting concerns: Tackling hardware, software, communication, usability and applicability issues all at once made it challenging to deploy a device that was fully functional since ongoing problems tended to escalate. It also posed difficulties in the implementation.

Obsolescence: Frequent updates to the hardware specifications required additional efforts from team members to redesign the electronics and the system software. Compounding this challenge, documentation and code comments did not always track the actual functionality, making it difficult for new team members to understand the software and hardware. In some cases, such discrepancies also led to flaws in the software and hardware. As an example, the original Gyro sensor became obsolete and was replaced by a newer version that was more power efficient. However, slight changes in the SPI initialization scheme caused the Gyro to never enter a sleep mode and increased quiescent draw 10x. This problem was only diagnosed after noting short battery lifetimes and questioning previous members of the project.

Management: Respondents noted challenges in maintaining alignment between the work of the team and the major goals of the project. They also found it difficult to maintain steady progress and ensure cross-team communication. Communication and coordination was hampered by the team's geographical distribution across remote labs and multiple campuses. The team met regularly (by video conference) and leveraged collaborative tools including Basecamp, Google Drive, and GitLab. Some team members noted that development could have been smoothed by ensuring best practices were shared and enforced across the entire team.

4.4 Deployment

When Amulet team respondents were specifically asked about the challenges they faced during the deployment studies, they reported five major issues.

Robustness: Respondents noted the Amulet would sometimes crash, that its mechanical design caused some discomfort, and that it was not waterproof; all of which limited the ability to deploy the device for extended periods. Sometimes the cases broke when study participants accidentally dropped the device; no standard material was enforced for the case - team members used what was available in home institutions (often brittle PLA). The capacitive touch slider was also a source of frustration, as it was a small PCB affixed via SMD pads to the main Amulet PCB, this slider sometimes came off.

Reliability: Respondents noted that the current application would sometimes be reset without warning. Because the Amulet does not have a hardware real-time clock, these issues resulted in some data points being logged with wrong date and time, leading to invalid data. Also, in one instance, the screen turned black, gray and then rendered back to normal. In another instance, data logging stopped working and the screen turned pale. Sometimes, the Amulet Operating System (OS) froze in the field and the device had to be reprogrammed in the lab by the investigator. While significant data gathering was enabled, some of these reliability issues (partially owing to the frequent updating of the platform and short deadlines) hampered more frequent data collection.

Ground truth validation: Depending on the study goals and outcomes, the Amulet team used different forms of ground-truth. In case of the stress studies, the team used responses to EMA prompts as the ground-truth stress level *in-the-field.* In the *Stress 3* study, the researchers were able to estimate the validity of the EMA responses by including two different five-point Likert-scale questions and testing the consistency in response to both the questions: (a) "Rate your stress over the last 10 minutes", ranging from "None (1)" to "Extremely High (5)", and, (b) "How do you feel right now?", ranging from "Really Bad (1)" to "Really Good (5)". The researchers assumed that both questions measure the same underlying trait, i.e., participant's stress, and hence, expected to see a negative correlation between the responses to these two questions across all participants. A Pearson correlation between these two items resulted in $r = -0.551$, $p < 0.001$. While a correlation coefficient makes sense intuitively, they also looked at the Cronbach's alpha measure for the two questions, and found the average $a = 0.711$. This result gave the researchers confidence in the validity and reliability of the EMA responses [23].

Fashion and wearability: Researchers noted that some participants complained about the aesthetics of the device. Participants commented that they did not like to wear the bracelet because it seemed like an older version of technology. Also, some of the early versions of the Amulet were more uncomfortable and bulky. As a result, some subjects mentioned that they did not wear the Amulet but preferred to place it inside their pockets.

Usability: For certain Amulet models, it was difficult to insert and remove the SD memory card, due to the tight dimensions of the card and of the bracelet, and the position of the card slot. Further, setting the date/time on the Amulet after a reset was tedious. Research staff had to read time.gov and manually set the time on the Amulets using the buttons and the touch slider. This process could have led to an error of ±1 second. Research staff also noted that there was a considerable clock drift in the Amulets (3 to 4 seconds a day). For multi-day studies, the time drift was corrected using post-processing scripts after the data collection was complete.

Energy: In one study, the battery level of one of the devices was fluctuating. In another, the Amulet device was not holding power, and would switch off unless permanently plugged into a power source. Because the Amulet does not have a hardware real-time clock, when the battery runs out and is recharged, the device resets its time to the time when it was initially programmed. This issue was a significant challenge, especially for apps whose functionality

relied on accurate timestamps. As a result, the time had to be either reset manually by users, which was potentially burdensome for subjects (and not a realistic option for older adults), or by researchers, which required subjects to return to the lab so the researchers could re-flash the Amulet. In retrospect it would have been useful and low design burden to have a built-in hardware real-time clock backed by a separate battery or super-capacitor.

5   LESSONS LEARNED

Based on the survey results, our interviews of the Amulet team, and reflection on our own experience as stakeholders in the five-year project, we derived a set of lessons learned concerning two categories: high-level management and hardware design specifics.

5.1   Project management

For a successful management of the technical components of the project, three recommendations were deemed essential: (1) a shared and reliable repository for maintaining and storing software and documentation, established early in the project; (2) standard conventions and practices for coding, implementation and especially documentation; and (3) a systematic development protocol involving periodic tests, as well as redesign and bug fixes when necessary. While these approaches are standard practice in industry, bringing this practice to academic projects is challenging, and difficult when combined with manuscript deadlines and reporting.

Development protocols and repositories.—We found need for the definition of a protocol for software management, including decisions about git branches, tags, rebasing, unit tests for new software and APIs (application programming interfaces), and test suites for new hardware. While industrial software-development procedures can lead to professional results, these practices are not always feasible for scientific projects conducted in academic settings. Nonetheless, the early adoption of consistent practices for coding standards and collaborative tools, such as shared repositories, becomes essential for efficient management of the project components and activities. To that end, standard practices for managing updates to software and hardware should be determined early and systematically adopted throughout the entire project. Importantly, as new students and staff join the project, they must be educated immediately on these specifics.

Documentation and accessibility.—Tutorials and training for new team members and external practitioners are important to ensure a better understanding of the system, to reduce the learning curve for novice stakeholders, and to incentivize broader adoption among stakeholders. This recommendation is particularly important when participants include non-tech savvy users. As a respondent commented "Make the documentation easy to read for novice users, including doctors, nurses, etc.".

Development process.—Full-stack development is challenging, especially in a multi-disciplinary mobile-health project. Very few team members had the required expertise to troubleshoot problems across the stack, meaning that these team members were in high demand when anything went wrong. Because of the moving hardware targets, often simple software bugs problems were blamed on hardware problems and left for later, leading to

delays in deployment or feature development. Careful time management, and allocation of resources for development and testing, are essential to ensure productivity as well as reliable device performance. It would have been helpful to settle on a single, reliable development environment from the beginning, to avoid issues with inter-operability. As one respondent commented on, early in the project "Anyone using MacPorts on MacOS could not install Brew to install the Amulet development environment. Working in VM's [Virtual Machines] is clunky and takes up large amounts of disk space. My development VM's would at times become corrupted and had to be recovered from backups" Eventually, the team codified the toolchain and wrote documentation, updating both for every release. Certain parts of the system were left as static libraries or binaries, so that developers did not need to start from scratch on every piece.

The challenges associated with distributed teams (located on multiple campuses) coupled with a high turnover of team members (inherent to academic projects in which student research assistants are involved), reinforce further the need for well-defined protocols for software and hardware management. Key motivation for PhD students (publishing papers) was not always aligned with good development practice. This tension was somewhat mitigated by team leaders impressing on junior researchers the importance of the device and platform to the medical and clinical community, and the future of healthcare.

Plan for obsolescence.—In any long-term development project the availability of components (chips, sensors, modules) will change; the Amulet team discovered (several times) that certain components were no longer available on the market, requiring alternate components to be selected and boards to be redesigned. Thus, a project should plan for hardware obsolescence from the beginning, by selecting components that are less likely to become quickly obsolete, and by selecting vendors with more stable support, documentation and maintenance records. As one respondent commented: "There was no way to know about many of the problems I described. With perfect foresight we could have stayed with the ANT radio, or chosen to use two chips for the BLE [Bluetooth Low Energy] radio, one for talking to smartphones, one for talking to sensors, but smartphones could talk to both so it seemed like buying the one chip that implemented both would be a relatively easy path. It turned out the vendors of the chip were being less than truthful about it's capabilities [sic]. This has become a problem in all kinds of purchasing of devices. Businesses are forcing customers to be beta and alpha testers for their products, especially on the software side." Unfortunately, this will likely be a problem in any research oriented hardware platform going forward as the pace of IC updates increase.

Project Scope.—While a general-purpose platform is versatile to support applications across domains, its implementation proved to be challenging to motivate, specify, and publish. By narrowing down the focus of a wearable health project to an application-specific domain, the development team may have been able to work with more focused goals and clearer activities in mind for both hardware and software design. As one respondent commented "Don't try to implement everything at once, work in small steps, vet your design choices, be ready to abandon things that do not work." Designing with a specific application in mind seems to be a safer approach, especially if it ensures enough flexibility to

accommodate additional features and yet meet the needs and requirements of other potential applications.

5.2 Design Specifics

Three recommendations regarding the specific design decisions are worth highlighting.

Aesthetics and robustness.—To improve the wearability of wearable devices, dedicated designers closely attuned to the mechanical demands of modern wearables should be closely involved throughout the project. User acceptance and comfort should be priority design requirements, to foster adoption and sustained engagement among end users, be those patients, study participants, or consumers. While technical specification (for both hardware and system) should be handled by a dedicated group and form-factor design by another, integrative workshops could allow them to unify design decisions and to address inherent trade-offs. The design team should also consider the durability, sturdiness, water-resistance, and robustness of the device to prevent damage from shocks and normal wear and tear during longitudinal deployment studies in the field with actual users.

Energy-efficiency.—Even with substantial advances in devising energy-efficient wearables, and the provision of dedicated tools for energy-efficiency analysis during app development, creating energy-efficient solutions still remains one of the top challenges in wearable health research and development. Stakeholders would benefit from a wider range of benchmark tests and deployment studies, to gain a better understanding about optimal sampling rates, estimated duration of battery life, and the mutual influences and trade-offs between context of use, actual usage, and stand-alone capabilities of a wearable device. Even now, energy-efficiency does not seem much of a concern in research wearable platforms or commercial, even though it is a major factor in adherence and population suitability.

Device specification.—Amulet should include a hardware-supported battery-backed real-time clock to facilitate fault-tolerant data collection. The Operating System should include internal logs to facilitate accountability and debugging when a unexpected behavior occurs. Other methods of in-situ debugging and crash reporting would have been useful, as long deployments tested the system in multiple ways.

As one respondent commented: "The Resource Profile needs to be updated and improved. Every smartwatch needs to have a real-time clock and it is very essential for having fully functional apps and smartwatch user experience in extended field studies."

6 SUMMARY AND REFLECTIONS

This paper reports the experiences of multiple stakeholders in designing, developing and evaluating an mHealth platform throughout a 5-year project. The main challenges our team faced involved system maintenance, integration, and management of cross-cutting concerns in the dynamic landscape of wearable technologies concerning both hardware and software. It also proved challenging to address inherent tradeoffs of multiple requirements, especially the tension between the goal of high usability for end users and programmers, robustness

and reliability for system implementation, and wearability and energy-efficiency for deployment studies. From our experience, our main recommendations include (1) clear, focused goals for the specification of hardware, (2) iterative design of hardware and software, (3) adoption of consistent development protocols and shared repositories, (4) regular and systematic testing, and (5) clear documentation, kept up- to-date. More specific recommendations concern design for aesthetics and robustness, and energy-efficient specification for similar wearable health platforms. We believe the experience presented in this paper will arm future computing and clinical researchers with valuable insight on all facets of a large scale mobile health project.

ACKNOWLEDGMENTS

REFERENCES

[1]. Baig Mirza Mansoor, GholamHosseini Hamid, Moqeem Aasia A., Mirza Farhaan, and Linden Maria. 2017. A Systematic Review of Wearable Patient Monitoring Systems - Current Challenges and Opportunities for Clinical Adoption. Journal of Medical Systems 41, 7 (19 6 2017), 115. 10.1007/s10916017-0760-1 [PubMed: 28631139]

[2]. Batsis John A., Boateng George G., Seo Lillian M., Petersen Curtis L., Fortuna Karen L., Wechsler Emily V., Peterson Ronald J., Cook Summer B., Pidgeon Dawna, Dokko Rachel S., Halter Ryan J., and Kotz David F.. 2019. Development and Usability Assessment of a Connected Resistance Exercise Band Application for Strength-Monitoring. World Academy of Science, Engineering and Technology 13, 5, 340–348. 10.5281/zenodo

[3]. Batsis John A., Zagaria Alexandra, Kotz David F., Bartels Stephen J., Boateng George G., Proctor Patrick O., Halter Ryan J., and Carpenter-Song Elizabeth A.. 2018. Usability evaluation for the Amulet wearable device in rural older adults with obesity. Gerontechnology 17, 3 (2018), 151–159. 10.4017/gt.2018.173.003.00 [PubMed: 30631251]

[4]. Batsis John A., Zagaria Alexandra B., Halter Ryan J., Boateng George G., Proctor Patrick, Bartels Stephen J., and Kotz David. 2019. Use of the Amulet in Behavioral Change for Geriatric Obesity Management. Journal of Digital Health 5 (6 2019). 10.1177/2055207619858564

[5]. Biopac Systems Inc. 2016. Biopac MP150. https://www.biopac.com/wp-content/uploads/MP150-Systems.pdf

[6]. Boateng George, Batsis John A, Halter Ryan, and Kotz David. 2017. ActivityAware: an app for real-time daily activity level monitoring on the Amulet wrist-worn device. In IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). IEEE, 431–435.

[7]. Boateng George, Batsis John A, Proctor Patrick, Halter Ryan, and Kotz David. 2018. GeriActive: Wearable app for monitoring and encouraging physical activity among older adults. In IEEE International Conference on Wearable and Implantable Body Sensor Networks (BSN). IEEE, 46–49.

[8]. Boateng George and Kotz David. 2016. StressAware: An app for realtime stress monitoring on the Amulet wearable platform. In IEEE MIT Undergraduate Research Technology Conference (URTC). IEEE, 1–4.

[9]. Boateng George G.. 2016. StressAware: App for Continuously Measuring and Monitoring StressLevels in Real Time on the Amulet Wearable Device. Technical Report TR2016–802. Dartmouth College, Computer Science, Hanover, NH. http://www.cs.dartmouth.edu/reports/TR2016-802.pdf

[10]. Boateng George G.. 2017. ActivityAware: Wearable System for RealTime Physical Activity Monitoring among theElderly. Technical Report TR2017–824. Dartmouth College, Computer Science, Hanover, NH. http://www.cs.dartmouth.edu/reports/TR2017-824.pdf

[11]. Motti Vivian Genaro and Caine Kelly. 2015. An overview of wearable applications for healthcare: requirements and challenges. In Adjunct Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the ACM International Symposium on Wearable Computers. ACM, 635–641.

[12]. Hardin Taylor, Scott Ryan, Proctor Patrick, Hester Josiah, Sorber Jacob, and Kotz David. 2018. Application memory isolation on ultra-low-power MCUs. In USENIX Annual Technical Conference (USENIX ATC). Boston, MA, 127–132.

[13]. Hester Josiah, Peters Travis, Yun Tianlong, Peterson Ronald, Skinner Joseph, Golla Bhargav, Storer Kevin, Hearndon Steven, Freeman Kevin, Lord Sarah, et al. 2016. Amulet: An energy-efficient, multi-application wearable platform. In Proceedings of the ACM Conference on Embedded Network Sensor Systems. ACM, 216–229.

[14]. Hillyard Peter, Luong Anh, Alemayehu Solomon Abrar Neal Patwari, Sundar Krishna, Farney Robert, Burch Jason, Porucznik Christina A., and Pollard Sarah Hatch. 2018. Comparing Respiratory Monitoring Performance of Commercial Wireless Devices. In arXiv preprint arXiv:1807.06767.

[15]. Iqbal Mohammed H, Aydin Abdullatif, Brunckhorst Oliver, Dasgupta Prokar, and Ahmed Kamran. 2016. A review of wearable technology in medicine. Journal of the Royal Society of Medicine 109, 10 (2016), 372–380. 10.1177/0141076816663560 arXiv:10.1177/0141076816663560 PMID: . [PubMed: 27729595]

[16]. Klugman Noah, Jacome Veronica, Clark Meghan, Podolsky Matthew, Pannuto Pat, Jackson Neal, Nassor Aley Soud, Wolfram Catherine, Callaway Duncan, Taneja Jay, and Dutta Prabal. 2018. Experience: Android Resists Liberation from Its Primary Use Case. In Proceedings of the Annual International Conference on Mobile Computing and Networking (MobiCom '18). ACM, 545–556. 10.1145/3241539.3241583

[17]. Kotz David, Gunter Carl A., Kumar Santosh, and Weiner Jonathan P.. 2016. Privacy and Security in Mobile Health - A Research Agenda. IEEE Computer 49, 6 (6 2016), 22–30. 10.1109/MC.2016.185

[18]. Lorincz Konrad, Chen Bor-rong, Challen Geoffrey Werner, Chowdhury Atanu Roy, Patel Shyamal, Bonato Paolo, and Welsh Matt. 2009. Mercury: A Wearable Sensor Network Platform for High-fidelity Motion Analysis. In Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems(SenSys '09). ACM, New York, NY, USA, 183–196. 10.1145/1644038.1644057

[19]. Lowens Byron, Motti Vivian, and Caine Kelly. 2015. Design recommendations to improve the user interaction with wrist worn devices. In IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops). IEEE, 562–567.

[20]. Lymberis Andreas and Dittmar Andre. 2007. Advanced wearable health systems and applications. IEEE Engineering in Medicine and Biology Magazine 26, 3 (2007), 29.

[21]. Mandalari Anna Maria, Lutu Andra, Custura Ana, Khatouni Ali Safari, Alay Özgü, Bagnulo Marcelo, Bajpai Vaibhav, Brunstrom Anna, Ott Jörg, Mellia Marco, and Fairhurst Gorry. 2018. Experience: Implications of Roaming in Europe. In Proceedings of the 24th Annual International

Conference on Mobile Computing and Networking (MobiCom '18). ACM, New York, NY, USA, 179–189. 10.1145/3241539.3241577

[22]. Mishra Varun, Lowens Byron, Lord Sarah, Caine Kelly, and Kotz David. 2017. Investigating contextual cues as indicators for EMA delivery. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the ACM International Symposium on Wearable Computers. ACM, 935–940.

[23]. Mishra Varun, Pope Gunnar, Lord Sarah, Lewia Stephanie, Lowens Byron, Caine Kelly, Sen Sougata, Halter Ryan, and Kotz David. 2018. The case for a commodity hardware solution for stress detection. In Proceedings of the ACM International Joint Conference and International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. ACM, 1717–1728.

[24]. Mitzner Tracy L, Boron Julie B, Fausset Cara Bailey, Adams Anne E, Charness Neil, Czaja Sara J, Dijkstra Katinka, Fisk Arthur D, Rogers Wendy A, and Sharit Joseph. 2010. Older adults talk technology: Technology usage and attitudes. Computers in human behavior 26, 6 (2010), 1710–1721. [PubMed: 20967133]

[25]. Motti Vivian Genaro and Caine Kelly. 2014. Human factors considerations in the design of wearable devices. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 58. SAGE Publications Sage CA: Los Angeles, CA, 1820–1824.

[26]. Motti Vivian Genaro and Caine Kelly. 2015. Micro interactions and multi dimensional graphical user interfaces in the design of wrist worn wearables. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 59. SAGE Publications Sage CA: Los Angeles, CA, 1712–1716.

[27]. Motti Vivian Genaro and Caine Kelly. 2015. Users' privacy concerns about wearables. In International Conference on Financial Cryptography and Data Security. Springer, 231–244.

[28]. Pantelopoulos Alexandros and Bourbakis Nikolaos. 2008. A survey on wearable biosensor systems for health monitoring. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 4887–4890.

[29]. Pantelopoulos Alexandros and Bourbakis Nikolaos G. 2010. A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Transactions on Systems, Man, and Cybernetics, Part C(Applications and Reviews) 40, 1 (2010), 1–12.

[30]. Paul Greig and Irvine James. 2014. Privacy implications of wearable health devices. In Proceedings of the International Conference on Security of Information and Networks. ACM, 117.

[31]. Petersen Curtis L., Wechsler Emily V., Halter Ryan J., Boateng George G., Proctor Patrick O., Kotz David F., Cook Summer B., and Batsis John A.. 2018. Detection and monitoring of repetitions using an mHealth-enabled resistance band. In IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). 21–23.

[32]. Piwek Lukasz, Ellis David A, Andrews Sally, and Joinson Adam. 2016. The rise of consumer health wearables: promises and barriers. PLoS Medicine 13, 2 (2016), e1001953.

[33]. Pope Gunnar, Mishra Varun, Lewia Stephanie, Lowens Byron, Kotz David, Lord Sarah, and Halter Ryan. 2018. An ultra-low resource wearable EDA sensor using wavelet compression. In IEEE International Conference on Wearable and Implantable Body Sensor Networks (BSN). IEEE, 193–196.

[34]. Sorber Jacob, Shin Minho, Peterson Ronald, Cornelius Cory, Mare Shrirang, Prasad Aarathi, Marois Zachary, Smithayer Emma, and Kotz David. 2012. An Amulet for trustworthy wearable mHealth. In Proceedings of the Workshop on Mobile Computing Systems & Applications. ACM, 7.

[35]. Sun Feng-Tso, Kuo Cynthia, Cheng Heng-Tze, Buthpitiya Senaka, Collins Patricia, and Griss Martin. 2010. Activity-aware mental stress detection using physiological sensors. In International conference on Mobile computing, applications, and services. Springer, 282–301.

CCS CONCEPTS

•Human-centered computing → Ubiquitous and mobile computing;•Applied computing → Life and medical sciences;•Computer systems organization → Embedded systems.
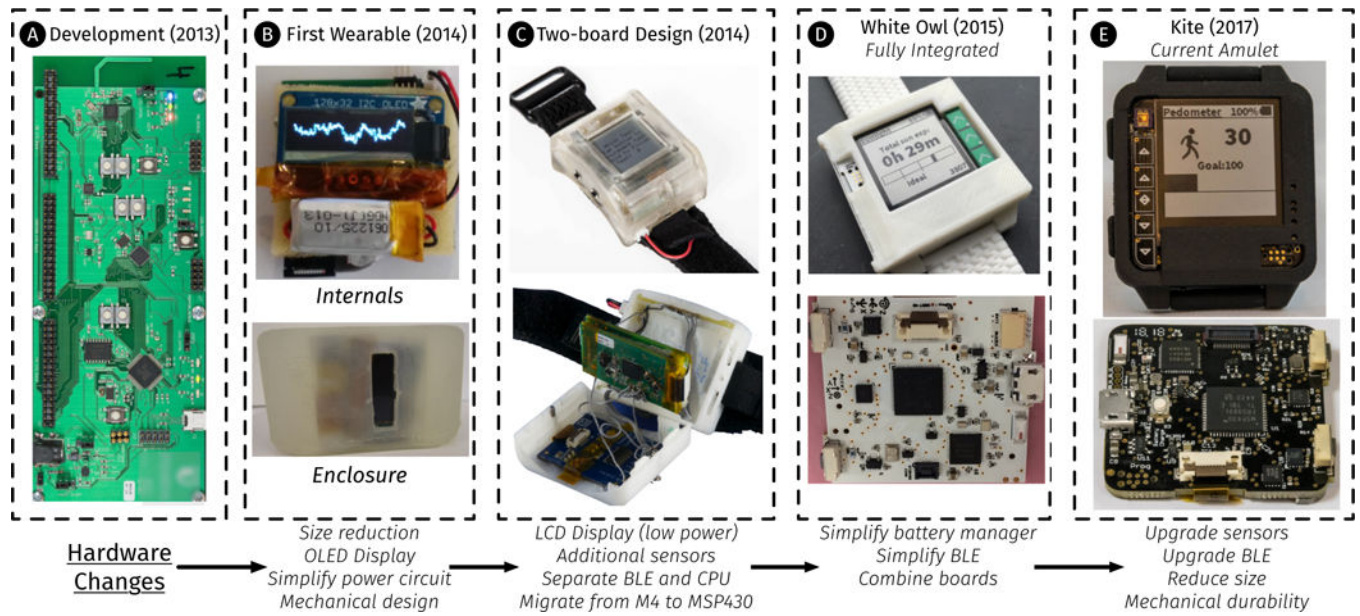
Figure 1:
Amulet wrist-worn mHealth platform

Figure 2:
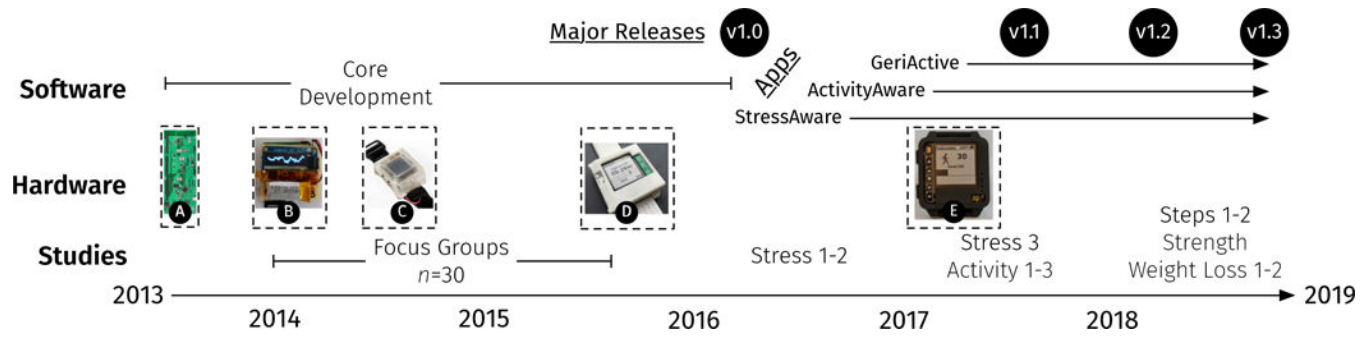Evolution of the Amulet hardware platform.

Figure 3:
Timeline of the Amulet project, studies, hardware and software development.

Table 1:

Eleven user studies were conducted with Amulet, to monitor stress levels, physical strength, number of steps, and physical activity level.

Study	N	Hours	EMA	Activity	HR	RRI	EDA	Steps	Force
				Data Gathered					
Stress 1	2	3	●	●	●	●			
Stress 2	10	100	●	●	●	●			
Stress 3	27	668	●	●	●	●	●		
Activity 1	29	29	●	●					
Activity 2	5	600		●					
Activity 3	7	504	●	●					
Steps 1	40	20	●					●	
Steps 2	12	576						●	
Strength	40	16	●						●
Weight 1	16	43,008		●				●	
Weight 2	16	32,256		●				●	