A Review of Commercial and Medical-Grade Physiological Monitoring Devices for Biofeedback-Assisted Quality of Life Improvement Studies

Pedro Nogueira [1,2] • Joana Urbano [1,2] • Luís Paulo Reis [1,3] • Henrique Lopes Cardoso [1,2] • Daniel Castro Silva [1,2] • Ana Paula Rocha [1,2] • Joaquim Gonçalves [4] • Brígida Mónica Faria [1,5]

[1] Artificial Intelligence and Computer Science Laboratory (LIACC), Porto, Portugal

[2] Faculty of Engineering of the University of Porto (FEUP), Porto, Portugal

[3] School of Engineering of the University of Minho (EEUM), Braga, Portugal

[4] EST/IPCA – Technology School/Polytechnic Institute of Cávado and Ave, Barcelos, Portugal

[5] ESS/P.Porto – Higher School of Health/Porto Polytechnic, Porto, Portugal

Abstract

With the rise in wearable technology and "*health culture*", we are seeing an increasing interest and affordances in studying how to not only prolong life expectancy but also in how to improve individuals' quality of life. On the one hand, this attempts to give meaning to the increasing life expectancy, as living above a certain threshold of pain and lack of autonomy or mobility is both degrading and unfair. On the other hand, it lowers the cost of continuous care, as individuals with high quality of life indexes tend to have lower hospital readmissions or secondary complications, not to mention higher physical and mental health. In this paper, we evaluate the current state of the art in physiological therapy (biofeedback) along with the existing medical grade and consumer grade hardware for physiological research. We provide a quick primer on the most commonly monitored physiologic metrics, as well as a brief discussion on the current state of the art in biofeedback-assisted medical applications. We then go on to present a comparative analysis between medical and consumer grade biofeedback devices and discuss the hardware specifications and potential practical applications of each consumer grade device in terms of functionality and adaptability for controlled (laboratory) and uncontrolled (field) studies. We end this article with some empirical observations based on our study so that readers might use take them into consideration when arranging a laboratory or real-world experience, thus avoiding costly time delays and material expenditures.

Keywords Psychophysiology · Quality of life · Biofeedback · Consumer grade hardware · Fitness tracking

Introduction

Originally starting with fitness watches and more recently with the advent of smartphone-powered fitness applications, recent years have seen an increase in the prevalence and complexity of wearable technology. This technological advancement has been made possible by the miniaturization of sensor technology and increased battery/circuit efficiency, which has been driven by an exponentially growing healthcare and "*health culture*" [1, 2].

One of the main reasons for the emergence of this "*health culture*" is the result of an increasing life expectancy of patients, which often isn't accompanied by their quality of life (QoL), thus leaving patients with great pain, restricted mobility and considerable adverse effects to their daily life and future health prospects. In other words, beyond prolonging life, it is essential to also increase patient's quality of life.

Quality of life is now considered an important aspect in clinical practice for patients with chronic illnesses, as someone with poor QoL indexes will likely suffer from a lack of mobility (due to, for example, joint pain), which in turn leads to low exercise and/or low mental health levels. These, in turn, slowly chip away at their health condition, aggravating their condition in a feedback loop and increase the overall risk of developing secondary health conditions.

Despite its overarching importance, the current methods to assess quality of life, automatic or semi-automatic, and its use in clinical decision support systems are still underexplored and there are virtually no applications in the market for this.

As a first step in developing a structured approach towards building such systems, in this paper we evaluate the existing state of the art in biometric analysis systems – both medical and consumer grade –, as well as existing biomedical studies on Quality of Life measurements and model building via biometric analysis. The paper is thus structured as follows: Section 2 describes the most common physiological metrics, as well as the current state of the art in biofeedback studies. Following this discussion, Section 3 expands on the previous topic by presenting the existing medical studies on Quality of Life improvements via biofeedback techniques. In Section 4, we discuss the existing medical and consumer grade devices in the market today. Section 5 concludes the paper by presenting an inter and intra device comparison, elaborating on the benefits and disadvantages of the uses of the presented devices for the purpose of measuring quality of life, and also offers some suggestions for their applicability to these studies.

Biometric data collection

The most objective way to evaluate quality of life is to monitor patients' physiological state over a relatively long period of time (ranging from weeks to, ideally, months). The most common way of doing this is through biofeedback techniques. Biofeedback itself was originally developed for medicinal purposes in the 1970s as a training procedure to overcome medical conditions, such as Attention deficit hyperactivity disorder (ADHD) [2]. However, in the last decade, it has re-emerged as a viable technology for use in ludic applications and some researchers have leveraged this by integrating biofeedback into interactive training or rehabilitation applications [3, 4]. This "sugar coating" makes the process more natural and less strenuous or boring for patients and can thus increase both the amount of data collected, as well as the efficiency of the program itself.

Despite this growing popularity, biofeedback apparatus is often medical-grade, and thus, expensive (ranging from $6000 to $15,000+ for a single device), meaning it is not readily available to the public or easy to apply to large studies. Several hardware manufacturers are attempting to provide inexpensive physiological input solutions that use brain signals (e.g., Emotiv Epoc[1], Neurosky Mindset[2], OCZ Neural Impulse Actuator[3]) and other physiological measures, such as skin conductance, oximetry, electromyography, respiration rates, and electrocardiography (e.g., BITalino[4]).

Throughout this section, we provide a review of the most current technology and research with regards to biofeedback for medical purposes. We begin however, with a short primer of physiological metrics prior to discussing these applications and the existing industry hardware.

A primer on physiological metrics

Electrodermal activity Electrodermal activity (EDA), usually measured in the form of skin conductance (SC), is a common measure of skin conductivity. EDA arises as a direct consequence of the activity of eccrine (sweat) glands. Some of these glands situated at specific locations (e.g., palms of the hands and feet soles) respond to psychological changes and thus EDA/SC measured at these sites reflects emotional changes as well as cognitive activity [4]. SC has been linearly correlated with arousal [3, 5–7] and extensively used as stress indicator [8], in emotion recognition [3, 8–10] and to explore correlations between gameplay dimensions [11, 12]. It is usually measured using two Ag/AgCL surface sensors snapped to two Velcro straps placed around the middle and index fingers of the non-dominant hand [7] but consumer-based hardware usually measures it on the wrist using metallic contacts, which reduces accuracy but is, understandably, a necessary trade-off.

Cardiovascular measures The cardiovascular system is composed by the set of organs that regulate the body's blood flow.

[1] https://www.emotiv.com/epoc/
[2] http://neurosky.com/
[3] https://www.ocz.com/us/
[4] http://bitalino.com/en/

Various metrics for its activity currently exist, among which some of the most popular ones are: blood pressure (BP), blood volume pulse (BVP) and heart rate (HR). Deriving from the BVP or HR, various secondary measures can be extracted, such as inter-beat interval (IBI) and heart rate variability (HRV). HR is usually correlated with arousal or physical exercise and can be easily differentiated using a combination of other sensors like, for example, skin conductivity [7]. HR, along with its derivate and HRV has also been suggested to distinguish between positive and negative emotional states (valence) [8, 9]. Heart rate is a very common measure for most sports watches and fitness bracelets and is usually measured via a photoplethysmogram (the volumetric measurement of an organ) using a pulse oximeter that illuminates the skin and measures changes in light absorption. For medical grade devices, the preferred method is to usually infer this from a raw ECG data stream (more precise) or the participants' BVP readings using a finger sensor (less precise).

Electromyography Electromyography (EMG) is a method for measuring the electrical potentials generated by contraction of muscles [7]. Facial EMG has been successfully used to distinguish valence in gameplay experiences [10]. In the former experiences, Hazlett describes the zygomaticus major (cheek) muscle as significantly more active during positive events and the corrugator supercilii (brow) muscle as more active in negatively-valenced events.

Respiration A Respiration sensor (RESP) measures the volume of air contained in an individual's lungs, as well as their breathing patterns. It is usually measured using a sensitive girth sensor stretched across the individual's chest [7]. It can be inferred indirectly through other methods such as an accelerometer or gyroscope but results are dependent on the uses and easily muddied by high physical movement so it is not advised on high precision scenarios.

Body temperature Body temperature (TMP) sensors are, in the vast majority of cases, highly-sensitive (able to measure changes up to a fraction of a Celsius degree) thermally sensitive resistors. These are quite cheap to create, consume a low energy output and are extremely small and easy to integrate within a watch or wearable technology. They represent one of the few cases where there is virtually no relevant difference between medical and consumer grade devices.

Medical biofeedback applications

As previously mentioned, originally biofeedback was designed to aid in medical therapy by helping patients overcome medical conditions or to perform patient monitoring/assessment [13, 14]. For example, a music therapy approach is presented by Dong et al. [14], where the users' negative emotional states are counterbalanced through music. In a similar approach, in [15] the authors presented a system to aid body balance rehabilitation by using simple audio frequencies to indicate correct posture. In related work, Huang et al. developed a neural motor rehabilitation biofeedback system for use in a virtual 3D world [16].

Due to biofeedback's easy integration with interactive and multimedia applications, various serious games have been designed to aid in the treatment of medical conditions. For example, a game was presented which targets the treatment of swallowing dysfunctions [13]. Riva et al. proposed a General Anxiety Disorder treatment that triggers changes in the game world based on the patient's heart rate and skin conductance [17]. A very similar biofeedback game ("Nevermind") for fear management based on players' heart rate readings was also designed [18].

Several approaches geared more towards self-improvement have also been proposed. For example, "Brainball" [19] and Bersak's proposed racing game [15] are relax-to-win indirect biofeedback games that introduce a competitive player-versus-player environment where the most relaxed player has a competitive advantage. While entertaining, the most interesting aspect of these games is their paradoxical design, because they combine two opposing concepts — relaxation and competitiveness. Naturally, in a competitive environment, players feel pressured to win. In turn, this hinders their ability to relax, which further puts them at a disadvantage and thus, under more pressure. This leads to a positive feedback cycle where the first player to achieve a competitive advantage tends to have increasingly higher odds of winning and thus benefits self-control and brain stimulation – a key factor in mental health and, therefore, quality of life.

Table 1 presents a comparative analysis of clinical biofeedback applications. For each work, we show the number of subjects used in the study (SS – sample size), the biofeedback type used (monitoring, direct biofeedback (DFB) or indirect biofeedback (IFB)), the adaptation mechanisms employed/driven by the physiologic data, area of application for the study, and finally the list of sensors used (BP = Blood Pressure, EMG = Electromyography, EEG = Electroencephalography, ACC = Accelerometer, PPG = Photoplethysmography, HR = Heart Rate, SC = Skin Conductivity).

Biofeedback for quality of life measurement and modelling

With increased life expectancy and overall survival rates in most diseases, as well as improved general living conditions, quality of life has become an important issue for most people. In healthcare, not only is the success of treating or managing a disease important but also the impact it has on the individual.

Table 1 Review of 10 medical and therapeutic applications of biofeedback techniques

Reference	SS	BF Type	Adaptations	Treatment	Sensors
Blanchard [20]	42	Monitoring	Thermal feedback	Elevated BP	BP
Bryant [13]	1	Monitoring	Muscle exercise regimen feedback	Swallowing Dysfunctions	EMG
Dong [14]	4	IBF	Musical excerpts	Music Therapy	EEG
Rocchi [21]	8	IBF	Audio Frequencies	Balance Control	ACC
Huang [16]	2	IBF	Musical and Visual Stimuli	Motor Rehabilitation	ACC, PPG
Stepp [22]	6	DBF	Control virtual fish	Swallowing Dysfunctions	EMG
Riva [18]	24	IBF	Virtual object placement and properties	General Anxiety Disorder	HR, SC
Reynolds [18]	NA	IBF	Audiovisual stimuli (game events)	Fear / Anxiety Disorders	HR
Hjelm [19]	NA	IBF	Ball movement / orientation	Relax to win	EEG
Bersak [15]	NA	IBF	Car acceleration	Relax to win	SC

Many people prefer death to a life without a level of quality they have become accustomed to. This makes it paramount to assess the quality of life of the individual, and how to improve it. With quality of life being an imprecise definition and dependent on the individual, it becomes necessary to develop models to understand and measure quality of life.

In [23], Wilson and Cleary describe a conceptual model for patient's quality of life improvement after medical procedures and/or therapies. They also address the fact that different patients respond differently to the same metrics, putting forth that this is a subjective metric that is not only influenced by the treatment itself, but also by how it is administered (e.g. patients being under the notion that they are constantly being monitored might make them feel more secure and thus less stressed about their own wellbeing, which in turn can lead to a more positive outcome regardless of treatment efficiency being the same with or without physiological monitoring).

Sprangers and Schwartz [24] present another QoL theoretical model to help patients cope with chronic or sudden life-threatening illnesses via 'response shift'. The model contemplates not only the patient's response to his health issue and subsequent treatment but also catalysts for it, antecedents, coping mechanisms and response shift (self-evaluation adaptation methods). The authors also point at the possibility of "a dynamic feedback loop aimed at maintaining or improving the perception of QoL" as future work that would drastically increase the effectiveness of QoL on these studies.

In [25], Felce and Perry discuss another theoretical QoL model based on a multidimensional analysis involving five dimensions: physical wellbeing, material wellbeing, social wellbeing, emotional wellbeing, and development and activity. It also allows/accounts for objective comparisons to be made between particular groups of individuals and what is normative.

Taillefer et al. [26] present a thorough review of theory-driven models of health-related QoL based on how well they tackle the methodological and conceptual problems in the field. The authors reviewed 68 models formulated from 1965 to 2001 based on a blind judge scheme that analyzed *a)* how sophisticated the models' conceptualization was; *b)* its definition of QoL; *c)* the distinction between factors that influence QoL; and *d)* the presence and/or usage of suitable instruments to measure QoL objectively. A grading scheme was introduced and interesting key points of this analysis revealed that 25% of authors did not define QoL per se and that while 78% of authors identified instruments to measure QoL, from the authors' analysis, none seemed to provide results based on actual measurements, much less based on large groups of low QoL patients on a continuous monitoring protocol.

In recent years, QoL has been defined as a measure of self-perception and its evaluation is performed through questionnaires. In the case of QoLRH (Quality of Life Related with Health) more than one questionnaire is normally used. There are dozens of questionnaires for QoL assessment that evaluate different QoL perspectives. The fact that these questionnaires are usually long and also the periodicity of administration (usually once a week or once a month, depending on the questionnaire) causes the respondents to give up on responding in a pondered manner, filling up the questionnaires in an careless manner, biasing the answers and, consequently, invalidating the measure. By way of example, the SF36V2 questionnaire has 36 questions and should be administered once a month and the EORTC-QLQC30 has 30 questions and should be administered once a week.

Thus, the use of questionnaires is incompatible with continuous monitoring, and biofeedback is profiled as an excellent alternative to the use of questionnaires mainly due to the almost non-existent intrusion into the individual's routines.

Most studies on quality of life have been focused on chronical diseases or others that somehow have a high impact on the patient's lifestyle, and sometimes requires the patient to make changes to life habits. In cancer in particular, QOL has been studies for more than three decades, with research works focused on particular cancer types and the impact on the individual quality of life [27–31], as well as more general research works by Aaronson [32] and Rehse [33].

Other diseases that cause social anxiety and discomfort, like constipation and fecal and/or urinary incontinence, have also been the target of several studies, to understand the impact the course of treatment has in the patient's life [34–40].

On a perhaps less evident or physical level, mental disorders and respective treatments also have a huge impact on the individual's perceived quality of life, and have also been studies focused on studying its nefarious effects on both the patients and their kin [41–43].

While most of these diseases affect a small percentage of the population, there are some conditions that either for a short or long-term affect most individuals. For example, Ahmedzai [44] has studied the impact of pain control on chronic pain patients and its influence on their quality of life, while Lasek and Chren [45] present a study on the impact of acne in adult quality of life, discussing mostly the mental trauma associated with it. These studies provide evidence that patients with serious medical conditions are not the only ones to benefit from the proposed monitoring systems and that there are suitable applications for a wider (perhaps mass) public. While the moral merits of such pursuit might not be obvious, the economic traction it entails might very well be a driving force behind further advances on this field.

Medical and consumer biofeedback devices

As seen above, biofeedback was originally developed for medical research applications, having only gained widespread traction in recent years. Thus, most of the existing state of the art research has been developed using medical grade devices, which essentially make no compromises in terms of accuracy but are somewhat lacking in practicality. Conversely, consumer grade biofeedback sensors, such as the ones present in modern fitness or lifestyle trackers, focus heavily on being "*everyday usable*" and trade sensor accuracy for other conveniences such as unobtrusiveness and battery life.

In this section, we start by succinctly describing the capabilities of the best-of-breed medical grade biofeedback hardware. We do so in order to contextualize the existing standard for sensor accuracy and hardware features so that we can establish a baseline for comparing consumer grade devices not only among themselves but also to the current state of the art. As our target study essentially requires research-level quality readings but also the amenities of everyday usability and autonomy, this is a necessary comparison. We then proceed to describe the most popular consumer grade devices before moving towards a comparative analysis and applicability considerations on the Discussion section.

Medical grade devices

Medical grade devices are designed with the goal of offering versatile but mostly highly advanced systems for physiological research. Most of them share the same technical specifications – as is required by strict medical guidelines – and hardware format – for compliance and competition purposes. The most popular solution on the market are the devices manufactured and sold by Mind Media BV, more specifically, the Nexus-10[5] series.

The Nexus-10 offers 8 input channels in total, with their configuration being customizable by the users for a particular set of sensors, depending on the study to be conducted. Mind Media offers a wide range of modalities that can be simultaneously measured. These include: electroencephalography (EEG), slow cortical potentials (SCP), electromyography (EMG), electrooculography (EOG), electrocardiography (ECG), blood flow via blood volume pulse (BVP), oximetry (O2), skin conductance (SC), respiration patterns (RSP), body temperature (TMP), accelerometer (ACC) and force sensors (FS).

From these, a wide range of secondary or processed variables can be extracted (e.g. heart rate, heart rate variability, breathing patterns, brainwave features, etc.) using the included software suite which is able to capture, collect and present the sensor data in real-time. It also allows users to configure custom dashboards and apply real-time filters to the data prior to logging them in several custom, text-based formats – including raw data outputs via Bluetooth to external applications.

On the practical side, the Nexus-10 (and all of its competitors) presents a simple but heavy data acquisition solution, roughly the size of a human hand (120x140x45 mm) and weighing around 500 g. While not too bulky, it is cumbersome and noticeable in every usage, especially if the patient is moving around or doing physical activity. Older versions of these devices were powered by 2 to 4 triple-A batteries, which lasted for roughly 10 h of capture with Bluetooth streaming active. Recent versions include a lithium-ion battery pack (8000 mAh lithium polymer) that reduces weight and should last for over 24 h, also making charging the device on the fly a possibility, which greatly helps in both lab and field studies.

Regarding sensor accuracy and sampling rates, the Nexus-10 series offers high quality medical grade connectors that isolate noise or artifacts when touching or pulling on them. The biggest source of noise is usually on the attaching surface sensors that, if not properly tightened/adhered, will provide erroneous readings. In terms of electrical interference, the carbon coating on the cables shields them and the active noise cancellation technology helps ameliorate the issues. Sampling rates are also the highest in the business with dual channel

[5] https://www.mindmedia.com/products/nexus-10-mkii/

inputs (ECG, EMG, EEG) recording data at 2048 Hz and single channel inputs (SC, BVP, TMP) and derived readings (e.g. HR) recording at 32 Hz.

In terms of sheer performance and autonomy, as we will see in the following section, the Nexus has essentially the edge over consumer grade devices. Where it loses ground is mainly in cost and its inability to offer a more compact package (most consumer grade devices are less than half of its weight and a fraction of the size). A single unit complete with the necessary cables and pre-gelled electrodes can easily cost over 8000€, which is enough to buy, on average, over 40 consumer grade devices and makes it highly inadequate for large or unsupervised studies.

Consumer grade devices

Unlike medical grade devices, designed for high-end research, consumer biofeedback is a recent trend that has been focused on ludic activities (e.g. biofeedback videogames), fitness tracking and lifestyle monitoring. Most of the available devices on the market have appeared in the last 4–5 years, having been made possible by 1) the miniaturization and mass production of sensor technology, mostly due to the advent of smartphones, and 2) by the increasing prevalence of "health cultures" and popular awareness of the importance of physical and mental well-being. In this section we evaluate eight consumer grade physiological recording devices, each from different manufacturers. We focus on the latest version of each of them and discuss them individually. In the following section, we present a comparative analysis between each of the consumer grade devices as well as how they stack up to medical grade devices. We will analyze the following dimensions:

- Price
- Available sensors
- Derived variables
- Existing API
- Software Suite
- Operating System Compatibility
- Communication protocols
- Battery Life

The first device on our list is the Feel Wristband from Feel. It retails for $199 and is designed to log emotion patterns throughout the day. How it achieves this is not described, as the algorithm is understandably proprietary, but there is an evident lack of scientific research backing this claim, which raises doubts as to its accuracy. From the available information on their website, it is evident they use skin conductivity, which has been shown to directly correlate with arousal [4] – one of the main emotional dimensions in Russell's circumplex model of affect. They also use a 3D accelerometer to track physical activity, and in all likelihood correct improper SC activations due to exercise; it is also likely they are computing the subject's heart rate for this. The biggest issue with this emotional detection process is that there is no discernable way of identifying valence – the second dimension in Russell's circumplex model – and without it, it is impossible to distinguish the emotion's positive or negative charge, only it's intensity. Secondly, an important distinction should be made between emotional states and emotions, as they are quite different. An emotional state is a coordinate in an N-dimensional emotional space (e.g. Russell's Arousal-Valence Cartesian space). It is an objective quantification but lacks context. An emotion on the other hand is not so well defined as it is more of a quickly shifting mood that relies heavily on context. For example, a high arousal, low valence emotional state can represent any emotion that has high energy/excitement and a very negative charge to it, such as for example, Terror, Stress or Anxiety. All of these emotions fall within the same emotional region. However, what emotion the emotional state translates to is highly dependent on context. This is a common issue with virtually all of the devices discussed in this section and manufacturers seem to be aware of this shortcoming so they often rely on the user to provide contextual information not only to do a black-box approach towards recognizing emotions in future dates but also as a fallback (and potentially psychological induction method) to validate their readings. The issue is not necessarily crippling to the usage of the device but lays itself to some doubts and thus, is not suitable for academic research.

The second device on our list is the Zenta Wristband[6] by Vinaya. It retails for $50 less than the Feel Wristband [7]($149) and offers the same general functionality with a significant amount of extra sensors. Overall, it presents SC, TMP, and O2 sensors. It also comes with an accelerometer and a microphone. From the included sensors it derives an impressing number of variables: HR, HRV, RSP, pulse transit time and pulse wave velocity, as well as a few higher-level ones, such as discrete emotions, calorie tracking and activity tracking. Both emotion prediction and calorie tracking require user input (overall mood and daily caloric intake, respectively).

Our third entry is the Microsoft Band 2[8], which is, by far, the most complete package on the market at the time of writing. It retails for the same price as the Feel Wristband ($199) but offers the following biometric sensors: SC, O2, and TMP. It also includes GPS, an ambient light sensor, a Gyroscope in addition to the standard accelerometer, an UV sensor, a barometer, and a microphone. It can compute the user's HR from the O2 sensor but it is not clear as to why RSP readings are not described on the technical sheet. It presents itself as the most

[6] https://www.indiegogo.com/projects/zenta-stress-emotion-management-on-your-wrist
[7] https://www.myfeel.co/
[8] https://www.microsoft.com/en-us/band

research-focused solution on the market and does not offer discrete emotion processing. It does however, act as a fitness tracker so it measures calorie intake/expenditure, sleep patterns and sleep quality analysis tools. It is also the only device on the market that offers a dedicated visualization and data processing suite (Microsoft Health), which is free of charge.

The fourth and fifth entries are the Basis Peak [9]and Jawbone UP3[10]. They retail for $199 and $129 but, oddly, the Basis Peak offers less sensors as it comes only with a O2 from which it derives BVP and HR and a gyroscope and accelerometer. The Jawbone on the other hand offers the same O2 sensor in conjunction with SC and TMP sensors from which it derives HR and RSP measures. It only includes an accelerometer but not a gyroscope. In terms of fitness tracking, both track caloric intake and expenditure, as well as sleep patterns, their quality and physical activity. None of them offer any data visualization platform or APIs to read the data in real-time.

The last three entries in our list are not necessarily wearable or dedicated biometric tracking platforms but aim to tackle the market through sheer volume or a low cost/personalization. The first of these is the Apple Watch 2[11], which retails starting at $269 and offers HR and RSP measurements though photoplethysmographic measurement and GPS tracking. The second platform is the Android counterpart to the Apple Watch, the Android Wear 2[12]. Contrary to the Apple Watch, the Android Wear is free and is composed of the SDK to develop wearable apps so it doesn't offer a dedicated hardware platform. As such, it is not possible to assess which sensors it offers. The biggest advantage to both of these platforms is that while they don't do most of the work for the user, they enable tech-savvy users and researchers to build their own applications from scratch and completely customize them, while also allowing them to access data in real-time and integrate with other existing apps on the marketplace to leverage their functionality. In short, they have the biggest potential. The main issue is the limited range of sensors on these wearable platforms, which can severely hinder the data collection ability. In this regards, the Android Wear has the edge as it can integrate with any compatible platform.

The final entry on our list, BITalino, addresses the main disadvantages seen on the consumer grade segment by offering a hybrid solution between these and medical grade devices. BITalino is an Arduino-based biometric solution designed for researchers and electronics/engineering hobbyists that want to build their own physiologic recording devices. It retails from $149 to $200 for the most complete pack, which includes all the necessary hardware, as well as cables and sensors for measuring ECG, EMG, EEC, SC and acceleration. At the moment it seems it doesn't have a temperature sensor available but these are readily available online and can be easily integrated into the platform. It offers a free and complete software suite (Open Signals) to visualize and process data offline or in real-time and can perform relatively complex signal and statistical analysis, as well as process the collected data to infer HR, HRV and RSP metrics. It is not as practical as most of the devices on the market and is not waterproof or water resistant as all of the devices discussed so far but, overall, it is possibly the most versatile and cost-efficient solution on the market.

Discussion

In this section we analyze each of the devices discussed on the two previous sections and compare them in terms of: (1) **Features**: "*How complete is the device and how much data can it produce?*" (2) **Signal Processing**: "*How much signal processing (e.g., filtering, noise reduction) does the device allow or require to extract meaningful information?*" (3) **Precision**: "*How accurately can the signal be interpreted for both raw and derived measures and how much noise interference is present?*" (4) **Sensor Reliability**: "*How likely is the sensor to fail and how much calibration does it require?*" (5) **Intrusiveness**: "*Would the required apparatus interfere with the daily lives of the candidates, potentially impairing the study or biasing it in any significant way?*" (6) **Practicality**: "*How long does the device's battery last and at what sampling rates?*" and (7) **Cost**: "*Based on the current retail prices, what would be the necessary budget to perform a medium-sized study (50-100 individuals)?*"

Table 2 presents a breakdown of the described devices in detail. It summarizes our previous discussion and allows us to quickly reference them for the purposes of this discussion.

Regarding the pure number of features (volume of data), most devices generate roughly the same outputs: HR, SC and some form of activity tracking via accelerometers or gyroscopes. The clear winners are the Microsoft Band 2 and BITalino platforms, which not only include all these outputs but also include a few not present on the competing devices. We observe the same pattern in terms or derived variables and available analysis tools as these are the only two devices that offer proper visualization, logging and analysis software for the collected data.

In terms of precision, it is difficult to evaluate how these devices will fare without a large, controlled field study with all of them but, in general, all devices gather data in the same way and feature similar sensors and sensor placements so, if properly used, results should not differ significantly. In terms of how it compares with medical-grade devices, it should be expected that more movement or electromagnetic interference

[9] https://sleeptrackers.io/basis-peak/
[10] https://jawbone.com/fitness-tracker/up3
[11] https://www.apple.com/lae/watch/
[12] https://www.android.com/wear/

be created given that these are lower grade devices, which can be alleviated with proper filtering. However, the greatest drawback is the fact that given that the sensors are placed in less-than-ideal body locations (e.g. SC should be measured in the index and middle fingers and most devices measure it on the wrist [4]), it makes the sensors prone to data collection failures, which in turn make the data stream incomplete and, ultimately, unusable. This is also aggravated by the longer-term nature of the study. As data has to be collected with as little flaws as possible over a long period of time, it is highly possible that patients will improperly readjust the sensors and not provide proper data.

This leads us into our fifth point, intrusiveness. While most devices, being worn on the wrist and being lightweight, are generally not very intrusive, none of them are waterproof and have been reported to be somewhat frail in terms of construction due to the sensor miniaturization. There have also been cases of user complaints regarding some minor discomfort when using them for prolonged times due to the blood flow restrictions or sensors digging into the wrist area. There was even a recall action on one of them due to the O2 sensor causing skin burns when used for long periods of time. This is also tied to practicality and another point of concern is the lack of substantial autonomy – all devices use a rechargeable LiPo battery that lasts anywhere between 24 and 96 h – and this implies either reducing the sampling rates to a bare minimum (e.g. 1 Hz or 5–10 min samples at 32 Hz per hour) or creating data collection pauses for charging the devices. Whether this is a pain point depends heavily on the focus of the study but should be carefully considered as it can put in question the validity of the study itself.

Our final analysis point concerns price and here all devices are generally within the same price range so there is not much discussion to be had. Clearly, the most attractively priced ones are the Jawbone UP3 and BITalino due to the sheer sensor/price ratio and the least attractive one is the Apple Watch 2. However, the differences in available sensors are basically nil when comparing between consumer and medical grade devices, so the main question should be whether medical grade devices' signal quality is needed and, if not, which devices are available for shipping on the market.

When considering all points, it seems clear that the most complete and versatile packages are the Microsoft Band 2 (MB2) and BITalino platforms. Where they differ is the fact that the MB2 does not require any significant assembly or coding to start collecting data and is much less intrusive than the BITalino. On the other hand, the BITalino offers a very complete package that, bar the higher signal precision, rivals medical grade devices. All in all, it is not possible to define the "best" device, as this will depend on the myriad of experimental conditions that are imposed by the study and need to be factored into the equation. Conditions such as sampling frequency, raw data output or processing capabilities, levels of acceptable intrusiveness or budget constraints, among others, can be decisive factors when choosing the device to use.

Given our own experiences with these devices we would, however, note that while on paper the MB2 is the most attractive from the price and number of sensors perspective, in reality, using these devices is riddled with constant issues on a day-to-day basis. Part of the issue is that these are only second-generation devices that have neither seen a wide adoption (and as a result, somewhat poor improvement cycles from the manufacturer) and that while they allow for access to the data they collect, this is not their focus and thus custom utility programs – and sometimes even hacks – are required to keep them running at the desired sampling frequencies and obtain the unprocessed data. The truth is that if budget is an issue but the time frame and man-hours allow for it, they can be a viable option but the lack of wide adoption and strict industry standards (usually present in medical grade devices) make the data

Table 2. Breakdown of currently existing consumer-grade physiological devices

Device Name	Price	Sensors / Raw Variables	Derived Variables	API	Suite	OS	Comm	Battery Life
Feel Wristband	$199	SC, HR, Acc	Emotions	Log	App	iOS/Android		48h, LiPo
Zenta Wristband	$149	SC, Tmp, HR, HRV, RSP, Acc, Noise	Emotions, calories, pulse wave/patterns	Log	App	iOS/Android		48h, LiPo
Microsoft Band 2	$199	SC, HR, Light, UV, Tmp, Gyro, Barometer, GPS, Noise	Calories, sleep tracking	Log	MS Health	iOS/Android		48h, LiPo
Basis Peak	$199	BVP, HR, Gyro	Calories, sleep tracking	Log	App	iOS/Android		96h, LiPo
Jawbone UP3	$129	SC, BVP, HR, Tmp, Acc	Calories, sleep tracking	Log	App	Desktop	Bluetooth	48h, LiPo
Apple Watch 2	$300	HR, GPS	Activity tracking	SDK	3rd Party Apps	iOS		24h, LiPo
Android Wear 2	$150+	NA	NA	SDK	3rd Party Apps	iOS/Android		NA
BITalino	$149+	ECG, EMG, EEG, SC, Acc	HRV, Signal filtering	SDK	Open Signals	Desktop iOS/ Android		NA

questionable from a scientific perspective, forcing the research team to substantiate it with comparative analyses and empirical studies to corroborate their findings.

Given all of this and the above, our opinion is that, for most scientific purposes, medical grade devices should be favored as they are superior in terms of data acquisition, quality, unforeseen failure rates and industry standards/certifications. If they are not an option due to budget or practical concerns the BITalino platform seems to be the best choice (assuming a technical team is present in-house) for the following situations: *a)* there is an option to do the study in a controlled or laboratory environment, *b)* there is a need to collect a wide range of physiological signals, *c)* signal quality should be medical grade but budget is a depriving factor and/or *d)* practicality can be somewhat compromised upon. For all other cases where practicality is a concern and there is no need to collect more complex physiological data such as EMG or EEG, the MB2 seems to be the best available option on the consumer device market.

Acknowledgements This article is a result of the project QVida+: Estimação Contínua de Qualidade de Vida para Auxílio Eficaz à Decisão Clínica, NORTE-01-0247-FEDER-003446, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). The authors also acknowledge to the strategic project LIACC (PEst-UID/CEC/00027/2013).

References

1. Swan, M., Health 2050: The realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. *Journal of Personalized Medicine* 2(3):93–118, 2012.
2. Lupton, D., Quantifying the body: monitoring and measuring health in the age of mHealth technologies. *Critical Public Health, Taylor & Francis* 23(4):393–403, 2013.
3. Chanel, G., Kronegg, J., Grandjean, D., Pun, T., Emotion Assessment : Arousal evaluation using EEG's and peripheral physiological signals in Proc. Int. Workshop on Multimedia Content Representation, Classification and Security, pp. 530–537, 2006.
4. Vinhas, V., Silva, D., Oliveira, E., Reis, L. Biometric Emotion Assessment and Feedback in an Immersive Digital Environment, *Soc. Robot.*, pp. 307–317, 2009.
5. Leon, E., Clarke, G., Callaghan, V., and Sepulveda, F., A user-independent real-time emotion recognition system for software agents in domestic environments. *Engineering Applications of Artificial Intelligence* 20(3):337–345, 2007.
6. Nogueira, P. A., Rodrigues, R., Oliveira, E., and Nacke, L. E., Modelling human emotion in interactive environments: Physiological ensemble and grounded approaches for synthetic agents. *Web Intelligence and Intelligent Agents.* 13(3):195–214, 2015. https://doi.org/10.3233/WEB-150321.
7. Stern, R. M., Ray, W. J., and Quigley, K. S., Psychophysiological recording. 2nd edition. New York: Oxford University Press, 2001.
8. Mandryk, R., and Atkins, M., A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies* 65(4): 329–347, 2007.
9. Hazlett, R., Measuring emotional valence during interactive Experiences : Boys at video game play, in Proc. of the SIGCHI conference on human factors in computing systems, pp. 1023–1026, 2006.
10. Pedersen, C., Togelius, J., Yannakakis, G. N., Modeling player experience for content creation, computational intelligence and AI in Games, 2(1), 121–133, 2009.
11. Nasoz, F., Lisetti, C. L., Alvarez, K, Finkelstein, N., Emotion recognition from physiological signals for user modeling of affect, In 3rd workshop on affective and attitude user Modelling, 2003.
12. Figueiredo R., Paiva, A., I want to slay that dragon - influencing choice in interactive storytelling, in Digital Interactive Storytelling, 2010.
13. Bryant, M. A. M., Biofeedback in the treatment of a selected dysphagic patient. *Dysphagia* 6(2):140–144, 1991.
14. Dong, Q., Li, Y., Hu, B., Liu, Q., Li, X., Liu, L. A solution on ubiquitous EEG-based biofeedback music therapy. In IEEE 5th International Conference on Pervasive Computing and Applications (ICPCA) (pp. 32–37). 2010.
15. Bersak, D., McDarby, G., Augenblick, N., McDarby, P., McDonnell, D., McDonald, B., Karkun, R. Intelligent biofeedback using an immersive competitive environment. 2001.
16. Huang, H., Ingalls, T., Olson, L., Ganley, K., Rikakis, T., He, J., Interactive multimodal biofeedback for task-oriented neural rehabilitation. In 27th annual international conference of the engineering in medicine and biology society (IEEE-EMBS), pp. 2547–2550. 2005.
17. Riva, G., Gaggioli, A., Pallavicini, F., Algeri, D., Gorini, A., Repetto, C., Ubiquitous health for the treatment of generalized anxiety disorders. In UbiComp '10. Copenhagen, Denmark. 2010.
18. Reynolds, E. Nevermind. URL: www.nevermindgame.com. 2013.
19. Hjelm, S. I., Browall, C. Brainball–using brain activity for cool competition. *NordiCHI*, 177–188. 2000.
20. Blanchard, E. B., Eisele, G., Vollmer, A., Payne, A., Gordon, M., Cornish, P., and Gilmore, L., Controlled evaluation of thermal biofeedback in treatment of elevated blood pressure in unmedicated mild hypertension. *Biofeedback and Self-Regulation* 21(2):167–190, 1996.
21. Rocchi, L., Benocci, M., Farella, E., Benini, L., Chiari, L. Validation of a wireless portable biofeedback system for balance control: Preliminary results. In IEEE Second International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth, pp. 254–257. 2008.
22. Stepp, C. E., Britton, D., Chang, C., Merati, A. L., Matsuoka, Y., Feasibility of game-based electromyographic biofeedback for dysphagia rehabilitation. In 5th International IEEE/EMBS Conference on Neural Engineering (NER), pp. 233–236. 2011.
23. Wilson, I. B., and Cleary, P. D., Linking clinical variables with health-related quality of LifeA conceptual model of patient outcomes. *JAMA.* 273(1):59–65, 1995. https://doi.org/10.1001/jama.1995.03520250075037.
24. Sprangers, M. A. G., and Schwartz, C. E., Integrating response shift into health-related quality of life research: a theoretical model. *Social Science & Medicine* 48(11):1507–1515, 1999. https://doi.org/10.1016/S0277-9536(99)00045-3.
25. Felce, D., and Perry, J., Quality of life: Its definition and measurement. *Research in Developmental Disabilities* 16(1):51–74, 1995. https://doi.org/10.1016/0891-4222(94)00028-8.
26. Taillefer, M. C., Dupuis, G., Roberge, M. A., and LeMay, S., Health-related quality of life models: Systematic review of the literature. *Social Indicators Research* 64(2):293–323, 2003. https://doi.org/10.1023/A:1024740307643.
27. Hollen, P. J., Gralla, R. J., Kris, M. G., and Potanovich, L. M., Quality of life assessment in individuals with lung cancer: Testing the lung Cancer symptom scale (LCSS). *European Journal of*

Cancer 29:S51–S58, 1993. https://doi.org/10.1016/S0959-8049(05)80262-X.

28. Shapiro, S. L., Lopez, A. M., Schwartz, G. E., Bootzin, R., Figueredo, A. J., Braden, C. J., and Kurker, S. F., Quality of life and breast cancer: Relationship to psychosocial variables. *J. Clin. Psychol.* 57(4):501–519, 2001. https://doi.org/10.1002/jclp.1026.
29. Carlson, L. E., Speca, M., Kamala, P. D., and Eileen, G., MindfulnessBased stress reduction in relation to quality of life, mood, symptoms of stress, and immune parameters in breast and prostate Cancer outpatients. *Psychosomatic Medicine* 65(4):571–581, 2003. https://doi.org/10.1097/01.PSY.0000074003.35911.41.
30. Kim, K. S., Lee, S. W., Choe, M. A., Yi, M. S., Choi, S., Kwon, S. H. Effects of abdominal breathing training using biofeedback on stress, immune response and quality of life in patients with a mastectomy for breast cancer, 35(7):1295–1303, 2005.
31. Faria, B. M., Gonçalves, J., Reis, L. P., Rocha, A., and Clinical Decision, A., Support system based on quality of life estimation. *Journal of Medical Systems* 39(10):308, 2015.
32. Aaronson, N. K., Meyerowitz, B. E., Bard, M., Bloom, J. R., Fawzy, F. I., Feldstein, M., Fink, D., Holland, J. C., Johnson, J. E., Lowman, J. S., Patterson, B., and Ware, J. E., Quality of life research in oncology: Past achievements and future priorities. *Cancer* 67(Suppl):839–843, 1991.
33. Rehse, B., and Pukrop, R., Effects of psychosocial interventions on quality of life in adult cancer patients: meta-analysis of 37 published controlled outcome studies. *Patient Education and Counseling* 50(2):179–186, 2003. https://doi.org/10.1016/S0738-3991(02)00149-0.
34. Bai, Y., Yuan, Z., Wang, W., Zhao, Y., Wang, H., and Wang, W., Quality of life for children with fecal incontinence after surgically corrected anorectal malformation. *Journal of Pediatric Surgery* 35(3):462–464, 2000. https://doi.org/10.1016/S0022-3468(00)90215-X.
35. Rothbarth, J., Bemelman, W. A., Meijerink, W. J., Stiggelbout, A. M., Zwinderman, A. H., Buyze-Westerweel, M. E., and Delemarre, J. B., What is the impact of fecal incontinence on quality of life? *Dis Colon rectum* 44(1):67–71, 2001. https://doi.org/10.1007/BF02234823.
36. Mason, H. J., Serrano-Ikkos, E., and Kamm, M. A., Psychological state and quality of life in patients having behavioral treatment (biofeedback) for intractable constipation. *The American Journal of Gastroenterology* 97(12):3154–3159, 2002.
37. Byrne, C. M., Pager, C. K., Rex, J., Roberts, R., and Solomon, M. J., Assessment of quality of life in the treatment of patients with neuropathic fecal incontinence. *Dis Colon rectum* 45(11):1431–1436, 2002. https://doi.org/10.1007/s10350-004-6444-1.
38. Burgio, K. L., Goode, P. S., Urban, D. A., Umlauf, M. G., Locher, J. L., Bueschen, A., and Redden, D. T., Preoperative biofeedback assisted behavioral training to decrease post-prostatectomy incontinence: A randomized, controlled trial. *The Journal of Urology* 175(1):196–201, 2006. https://doi.org/10.1016/S0022-5347(05)00047-9.
39. Byrne, C. M., Solomon, M. J., Young, J. M., Rex, J., and Merlino, C. L., Biofeedback for Fecal Incontinence: Short-Term Outcomes of 513 Consecutive Patients and Predictors of Successful Treatment. *Dis Colon rectum* 50(4):417–427, 2007. https://doi.org/10.1007/s10350-006-0846-1.
40. Rivalta, M., Sighinolfi, M. C., Micali, S., Stefani, S., and Bianchi, G., Sexual Function and Quality of Life in Women with Urinary Incontinence Treated by a Complete Pelvic Floor Rehabilitation Program (Biofeedback, Functional Electrical Stimulation, Pelvic Floor Muscles Exercises, and Vaginal Cones). *The Journal of Sexual Medicine* 7(3):1200–1208, 2010. https://doi.org/10.1111/j.1743-6109.2009.01676.x.
41. Spitzer, R. L., Kroenke, K., Linzer, M., Hahn, S. R., Williams, J. B. W., deGruy, F. V., Brody, D., and Davies, M., Health-related quality of life in primary care patients with mental DisordersResults from the PRIME-MD 1000 study. *JAMA.* 274(19):1511–1517, 1995. https://doi.org/10.1001/jama.1995.03530190025030.
42. Awad, A. G., Quality-of-life issues in medicated schizophrenic patients. In: Shriqui, C. L., Nasrallah, H. A. (Eds), Contemporary issues in the treatment of schizophrenia. Washington DC: American Psychiatric Press, 1995, 735–747.
43. Brod, M., Stewart, A. L., and Sands, L., Conceptualization of quality of life in dementia. *Journal of Mental Health and Aging* 5:7–19, 1999.
44. Ahmedzai, S., Recent clinical trials of pain control: Impact on quality of life. *European Journal of Cancer* 31A:S2–S7, 1995.
45. Lasek, R. J., and Chren, M., Acne vulgaris and the quality of life of adult dermatology patients. *Arch Dermatol.* 134(4):454–458, 1998. https://doi.org/10.1001/archderm.134.4.454.