RADAR-base: Major Depressive Disorder and Epilepsy Case Studies

Callum L Stewart
Institute of Psychiatry, Psychology & Neuroscience (IOPPN) King's College London
callum.stewart@kcl.ac.uk

Richard JB Dobson
Institute of Psychiatry, Psychology & Neuroscience (IOPPN) King's College London
richard.j.dobson@kcl.ac.uk

Zulqarnain Rashid
IOPPN, King's College London
zulqarnain.rashid@kcl.ac.uk

Amos A Folarin
IOPPN, King's College London
amos.folarin@kcl.ac.uk

Yatharth Ranjan
IOPPN, King's College London
yatharth.ranjan@kcl.ac.uk

The RADAR-CNS Consortium
https://www.radar-cns.org/

Shaoxiong Sun
IOPPN, King's College London
shaoxiong.sun@kcl.ac.uk

UbiComp/ISWC '18 Adjunct, October 8-12, 2018, Singapore, Singapore

ACM ISBN 978-1-4503-5966-5/18/10...$15.00
https://doi.org/10.1145/3267305.3267540

Abstract
Emerging mobile health (mHealth) and eHealth technology could provide opportunities for remote monitoring and interventions for people with mental health and neurological disorders. RADAR-base is a modern mHealth data collection platform built around Confluent and Apache Kafka. Here we report progress on studies into two brain disorders: major depressive disorder and epilepsy. For depression an ambulatory study is being conducted with patients recruited to three sites and for epilepsy an in-hospital study is being carried out at two sites. Initial results show smartphones and wearable devices have potential to improve care for patients with depression and epilepsy.

Author Keywords
mHealth; mobile context sensing; wearable sensors; data collection platform; mental health

ACM Classification Keywords
H.5.m [Human-centered computing (HCC)]: Ubiquitous and mobile computing.

Introduction
There has been an enormous increase in the capability to monitor individuals via smartphones and wearable devices during the last decade, with a growing range of parameters

offered by such technologies for continuous measurement [17].

The € 22 million Innovative Medicines Initiative (IMI2) Remote Assessment of Disease and Relapse - Central Nervous System (RADAR-CNS) is a major research programme aimed at developing novel methods and infrastructure for measuring Major Depressive Disorder (MDD), Epilepsy (EPI), and Multiple Sclerosis (MS) using wearable devices and smartphone technology [5].

The RADAR-base platform [1] is developed to support the three initial goals of RADAR-CNS, but importantly it has been developed such that it can easily be adapted for the needs of other mental and physical health disorders. The platform enables study design and set up, active and passive remote data collection. It provides secure data transmission and scalable solutions for data storage, management and access. This paper will focus on the MDD (RADAR-MDD) and EPI (RADAR-EPI) studies which use remote and in-hospital deployments of the RADAR-base platform respectively.

Major depressive disorder, sometimes called "clinical depression" or "depression", can be triggered by a life event, or result from stress, or happen without a specific cause. It is the most severe form of depression where people exhibit a sense of hopelessness and despair along with low mood and negative thoughts. This can affect the way people eat, sleep, feel about themselves, and think about things. Without treatment, the symptoms can last for weeks, months, or even years.

The RADAR-base platform has been deployed centrally to collect active (questionnaires) and passively generated (wearable and smartphone sensor) data remotely for patients recruited to 3 sites of MDD study. The sites include King's College Hospital (KCH) London, Centro de Investigacion Biomedica en Red (CIBER) Barcelona and VU University Medical Center Netherland. The objective being to collect regular self reported symptoms and metrics such as sleep and ambulatory behaviour. High resolution data is being collected over a period of up to two years for each participant.

Epilepsy is a neurological condition characterised by a person's tendency to have epileptic seizures. The global prevalence of epilepsy is between 4-10 per 1000 people. Those with epilepsy have a reduced life expectancy; people with symptomatic epilepsy have a life expectancy 18 years shorter [6]. Our hypothesis is that consumer type wearable devices have the potential to provide continuous seizure detection which may enable more informed use of anti-epileptic drugs, generating a more objective view of a person's condition.

Though current hospital observational systems (Video/ EEG/ ECG) are used in home monitoring they are not practical for long term epilepsy seizure detection within home based settings. We are using the RADAR-base platform to explore the feasibility of three wearable devices to detect seizures in an ambulatory settings. Data is being collected for a maximum of 14 days per patient.

These two studies expose the versatility of the RADAR-base platform and generate data with very different complexity, volume, velocity and durations.

Related Work

A number of relevant studies and mHealth platforms for remote monitoring in mental health are discussed here [17].

HORYZONS is a web based interface and feedback system to study people with first episode psychosis (FEP), a

one month pilot study with 20 participants was conducted [2]. The study aim was to provide an Internet-based intervention to young people with psychosis, to provide cost-effective long-term treatment to sustain the benefits of early intervention. The majority (75%) reported that they had a positive and constructive experience using the system, however this was a short term pilot with limited participants focused on young population with FEP.

Another study involves the naturalistic follow up of responders from the study entitled "Integrated biological markers for the prediction of treatment response in depression", or the **CBN-Well** study. In this study, participants who are currently responding to an oral antidepressant treatment regimen and/or therapeutic intervention were monitored over a minimum period of 13 months, providing an important opportunity to discover near-term biomarkers of relapse [12].

OBSERVEMDD a prospective, multicenter, longitudinal, single-cohort, observational study with MDD participants was performed using accelerometers and smartphone delivered questionnaires [18]. MDD patients who responded to, and continue to respond to an oral antidepressant treatment regimen were selected. The study consisted of 2 parts: a screening phase of up to 2 weeks, and an observational phase of variable duration. A total of 350 participants were recruited.

The RADAR-CNS programme advances the field in a number of ways. Other studies to date have made little or limited use of multi-parametric remote monitoring (RMT) by combining different sensors to detect signatures helpful for predicting outcomes in MDD. RADAR-MDD will take advantage of the combination of multiple sensor types along with remote data collection from a clinical population.

Detection of seizures using non-EEG wearable devices has been reasonably well studied over the last decade [7]. However, performance of the proposed models has often been unsatisfactory, particularly in terms of specificity, although some studies do show some promising results for seizures with a large motor component [3]. Additionally, few studies have been conducted outside of an in-patient environment, so the performance of these models in the real-world is unknown. To address the accuracy issues of models that only use a single sensor type, usually an accelerometer, there has been a movement within the field towards detection using multiple modalities[7]. A few studies have used using multiple sensors. Poh et al. used electrodermal activity (EDA) and an accelerometer, and showed increased GTCS detection performance when using both as opposed to only acceleration in 7 patients[16]. Heldberg et al. also used EDA and an accelerometer, looking at both convulsive and non-convulsive seizures in 8 patients [11]. Other studies have looked at the combination of acceleration and ECG-derived cardiac features[19, 9]. The use of multiple sensors does not always uniformly lead to better performance; Milosevic et al. report improved seizure detection but lower specificity when using both accelerometers and electromyography[15]. Finally, through the RADAR-base platform we have developed a well engineered open source platform with highly generalizable capabilities.

Methods

Remote Data Collection for Major Depression

RADAR-MDD, the major depression clinical substudy of RADAR-CNS, makes use of a range of data collection instruments as discussed below.

Passive RMT (pRMT) app

The passive application runs in the background, requiring minimal or no input from participants. Data is collected from smartphone "sensors" corresponding to a range of cate-

gories considered putatively relevant to the study, including (i) movement sensors: acceleration, gyration, and steps, and obfuscated relative GPS location; (ii) social characteristics: call duration, a log of SMS communications, contact list, and nearby Bluetooth handshakes; (iii) environmental sensors: ambient light, battery level, magnetic field, and weather conditions; (iv) user interaction with other applications and their phone; and (vi) keystrokes are collected in a subsample. All the collected data is pseudonomised, for example by hashing contacts names and phone numbers, and by using an unknown offset to obfuscate location.

Wearable Sensors

The Fitbit Charge 2 was selected to be worn by participants in RADAR-MDD for the duration of the study, providing metrics derived from the watch accelerometer and photoplethysmography (PPG). These data are processed on the device by vendor algorithms to provide information on heart rate, movement, daytime and sedentary activity, physical exercise, step count, and sleep patterns and efficiency. Data is collected into the RADAR-base platform from the Fitbit Web API, using the 3rd Party Data Integration service.

Active RMT (aRMT) app

Variation in the depression symptoms are measured via the 8-item Patient Health Questionnaire (PHQ8) [13] every 2 weeks throughout the course of follow-up. Variation in self-esteem is measured using the Rosenberg Self-Esteem Scale (RSES) [10]. The RSES is a widely-used 10-item self-reported questionnaire used to quantify self-esteem along a continuum and is administered alongside the PHQ8 every 2 weeks. As with the PHQ8 and RSES, every 2-weeks participants are asked to complete a speech task. This requires participants to read aloud, in a quiet area, some excerpts from "The North Wind and the Sun", which has been shown to be phonetically balanced across all three languages [20]. The excerpts are offered on a random schedule to prevent rehearsal and fluency and preserve prosodic features. In addition to this, participants are asked to respond to the following question: "Can you describe something you are looking forward to this week?". The aRMT app also delivers an Experience Sampling Method (ESM) schedule, designed to collect brief, in-the-moment assessments relating to several domains of interest: mood, stress, sociability, activity and sleep. Participants will receive a series of questions intended to reflect their current state (such as "right now, I feel content"), with 7-point Likert scale answer options (0=Not at all, 7 = Very much). The ESM schedule consists of approximately 44 items, taking up to 3-minutes to complete, delivered 9 random times per day within 90-minute blocks starting from 08.30 and ending at 22.00 for 6 consecutive days every 6 weeks.

THINC-IT app

THINC-it is a third party app used to assess cognitive function both objectively and subjectively, validated for detecting cognitive dysfunction in patients with MDD[14]. It incorporates four game-like digital assays, variants of widely-used cognitive assessments and a 5-item questionnaire assessing perceived deficits in memory, concentration, and attention over the previous week.

In-Hospital Data Collection for Epilepsy seizure detection

Hospital in-patient participants are recruited for the RADAR-EPI substudy of RADAR-CNS study as part of an otherwise typical stay at the Clinical Neurophysiology Department at Kings' College Hospital (KCH), London, UK or the Epilepsy Center at the University Hospital of Freiburg, Germany. Patients are monitored by a video-EEG and seizures are annotated by clinicians as part of the routine clinical assessment of their seizures. This provides ready-made source of

gold-standard labels for use in developing wearable-based seizure detection methods. In parallel, each patient wears 1 to 3 of the study wearable devices; the Empatica E4 wristband, Faros 180, Biovotion VSM1, or an offline IMEC device. The IMEC device records data offline, which is routinely transferred to the RADAR-base storage server. The other devices send data to the RADAR-base platform via a Bluetooth-paired android device through the passive RMT app.

Passive RMT (pRMT) app
The RADAR-base passive app (pRMT) has the capability to quickly integrate data sources (via pRMT plugins) such as wearable devices. The Empatica E4, Faros 180, and Biovotion VSM devices have been integrated for the EPI study. Each device was selected because of its ability to monitor physiologically relevant parameters. Acceleration, EDA, and heart rate. Cardiac features are measured either by PPG in the Empatica and Biovotion devices, or by ECG in the Faros device. The IMEC, although not integrated into the pRMT due to unavailability of a software development kit but has similar sensors to the Faros. Raw data is collected directly over Bluetooth (in comparison to the Ftibit where data is retrieved from the vendor data warehouse).

Study Population

As part of the RADAR-CNS programme RADAR-base is deployed to carry out RADAR-CNS studies at 8 sites across Europe, with the goal of enrolling MS (n=640), MDD (n=500) and EPI (n=200) participants.

Current Status of the MDD and EPI Studies
At present there are 66 enrolled patients in the MDD study at KCH. So far there has been a total of 127 patients enrolled in the EPI study across KCH and Freiburg.

Statistical and Analysis Plan

Preliminary analysis of the MDD dataset will investigate correlations between the PHQ-8 scores and basic aggregated features, obtained from the recording biosensors, which should be representative of behaviours associated with depression. The PHQ-8 questionnaire is taken every 2 weeks, and so the outcome is at a much lower frequency than the raw signals. Simple proxies for sleep, activity, sociability, cognition, and ambulation will be used to classify current depression, where current depression is determined by a PHQ-8 score ≥ 10. The ability of those features both to detect depressive periods between subjects and to monitor the progression of depressive symptoms within individuals will be explored. Rarer clinical relapse will also be reported for the cohort over the 2 year data collection period providing more definitive outcome measure where present. It may be necessary to design features in such a way that they are able to deal with missing data, or else use a model that is able to use missingness informatively [4].

The initial analysis of the epilepsy data requires a different approach. Seizures are typically short and sparse, so the primary challenge is to be able to detect the relatively short periods of ictal activity between the much more common segments of interictal time, while keeping the false positive rate at an acceptable level. Initial focus will be on the offline detection of generalised and focal seizures with a motor component, particularly those with tonic or clonic movements, using a combination of the available signals. The combination of the different signal modalities should improve the estimation accuracy of a single relevant parameter, and also allow the analysis using multiple physiologically-relevant parameters, enabling a panoramic view of the patient's status.

Because a few participants have had a large number of

seizures (n > 15) during their in-patient stay, there is an opportunity to measure the performance of an individualised seizure detection algorithm as successive seizures are added to the model. This will have application to future ambulatory studies, in which it will be important to know how much data is required for an adequately accurate model. Firstly, we will follow a analytical pipeline similar to those in prior seizure detection work, extracting features from the EDA, accelerometer, and heart rate signals, and classifying the ictal period of focal motor seizures using a support vector machine. Although the specificity of previous work has been too low, the larger sample size of the EPI study may help improve accuracy of similar models. Additionally, we will try and determine feature importance with the intention of elucidating performance gain from including additional signal modalities. Subsequently, we will investigate the feasibility of using deep learning techniques which may provide better generalization. Given the relative sparsity of ictal data, it may be necessary to use unsupervised neural networks to extract features, or to use transfer learning from the activity recognition domain.

Results and Discussion

A tonic EDA response during the post-ictal period has been noted elsewhere[8], and often occurs within the RADAR-EPI dataset. An example is given in Figure 1, showing an Empatica E4 recording of acceleration and EDA over a night-time 5-hour period. The convulsive seizure at 05:05 is followed by a large increase in skin conductance, with a peak at 05:10. There are other tonic peaks in the EDA, but they do not coincide with a seizure-like accelerometer trace. Equally, there is not evidence accelerometer traces with repetitive or otherwise confusable characteristics in the inter-ictal period being succeeded by an EDA response. Although not totally consistent across all participants and all seizures, it is a general pattern that illustrates the potential to use multiple modalities for increased specificity.

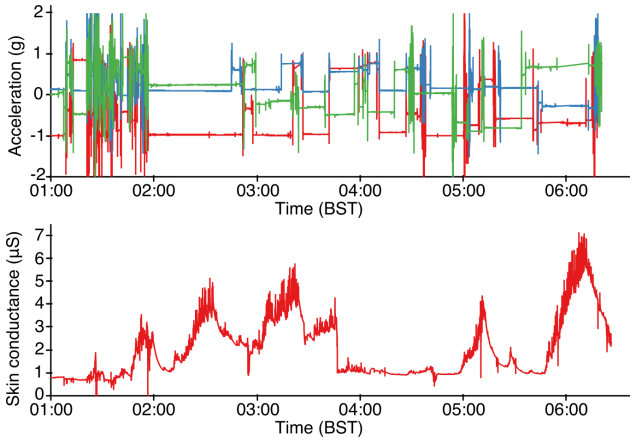

Figure 1: Data stream from a patient wearing an Empatica E4 during a night. The patient had a focal motor seizure at 05:05 (BST), corresponding to a burst of movement in the accelerometer (top), and subsequently followed by a peak in EDA (bottom). Other movements and peaks in EDA during the interictal periods do not follow the same pattern.

The preliminary data from the MDD study shows a range of depressive symptoms, with a mean PHQ-8 score of 10.4 and standard deviation of 6.2 in the 76 PHQ-8 questionnaires so far recorded. Five participants have had a depressive episode, progressing from a PHQ-8 score < 10, no depression, to a score >= 10, current depression, in the following questionnaires. Of those, one returned to a 'no depression' state after a week. There is, therefore, already a small amount of intra-individual variation recorded, although longitudinal effects should become clearer as the follow-up data collection period continues.

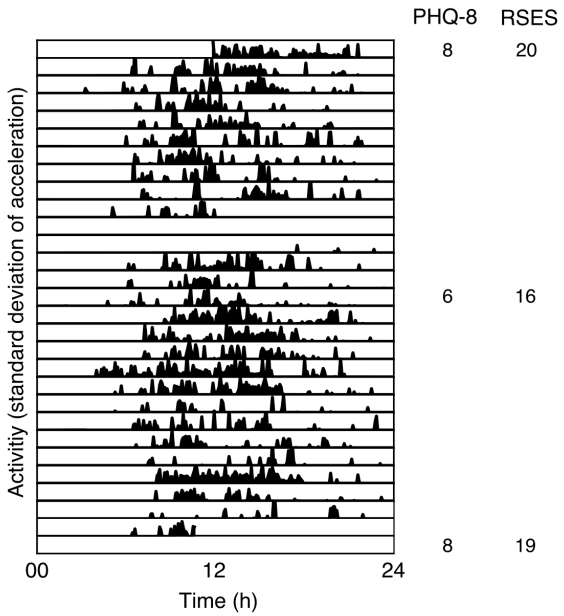

Figure 2: Activity of a participant as measured by the standard deviation of their phone's acceleration with corresponding PHQ-8 and RSES scores collected on the first day and every fortnight thereafter. Each row corresponds to a day. The questionnaire scores suggest the participant is not in a depressive mood (PHQ-8 scores < 10) and has a normal level of self-esteem (RSES scores between 15-25). There is missing data on days 10-12.

Missing values may prove a challenge for the MDD analysis. Firstly, due to technical challenges associated with a project of this magnitude. Secondly, and more commonly, through participant non-adherence and differing levels of engagement with the study applications and their phone in general. Even in patients with high adherence, there are likely to be times during which data is not collected. Figure 2 shows the first month of accelerometer data from a participant, alongside PHQ-8 and RSES responses. Although overall adherence is high for this participant, there is still a 48 hour gap during which no data is available. Disentangling missingness due to technical issues and missingness due to the participant non-adherence, and then directly incorporating that information into a model may be important, because depressive symptoms may affect adherence.

Acknowledgements

This work has received support from the EU/EFPIA Innovative Medicines Initiative Joint Undertaking 2 (RADAR-CNS grant No 115902) www.imi.europa.eu. This communication reflects the views of the RADAR-CNS consortium and neither IMI nor the European Union and EFPIA are liable for any use that may be made of the information contained herein. We would like to acknowledge The Hyve and RADAR-CNS Consortium (http://www.radar-cns.org/partners) for their support. Backend Infrastructure facilities were provided by King's College London Rosalind. The Authors receive funding support from the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London.

References

[1] 2018. RADAR-base. https://radar-base.org/. (2018).

[2] M. Alvarez-Jimenez, S. Bendall, R. Lederman, G. Wadley, G. Chinnery, S. Vargas, M. Larkin, E. Killackey, P.D. McGorry, and J.F. Gleeson. 2013. On the HORYZON: Moderated online social therapy for long-term recovery in first episode psychosis. *Schizophrenia Research* 143, 1 (2013), 143 – 149. DOI: http://dx.doi.org/https://doi.org/10.1016/j.schres.2012.10.009

[3] Sándor Beniczky, Isa Conradsen, Oliver Henning, Martin Fabricius, and Peter Wolf. 2018. Automated real-time detection of tonic-clonic seizures using a wearable EMG device. *Neurology* (2018). DOI: http://dx.doi.org/10.1212/WNL.0000000000004893

[4] Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. 2018. Recurrent neural networks for multivariate time series with missing values. *Scientific reports* 8, 1 (2018), 6085.

[5] RADAR-CNS consortium. 2016. RADAR-CNS: Remote Assessment of Disease and Relapse in Central Nervous System Disorders. https://www.radar-cns.org/. (2016).

[6] Hanneke M. de Boer, Marco Mula, and Josemir W Sander. 2008. The global burden and stigma of epilepsy. *Epilepsy and Behavior* 12, 4 (2008), 540 – 546. DOI: http://dx.doi.org/https://doi.org/10.1016/j.yebeh.2007.12.019 Current Views on Epilepsy and Behavior.

[7] Anouk Van de Vel, Kris Cuppens, Bert Bonroy, Milica Milosevic, Katrien Jansen, Sabine Van Huffel, Bart Vanrumste, Patrick Cras, Lieven Lagae, and Berten Ceulemans. 2016. Non-EEG seizure detection systems and potential SUDEP prevention: State of the art: Review and update. *Seizure* (2016).

[8] Onorati Francesco, Regalia Giulia, Caborni Chiara, Migliorini Matteo, Bender Daniel, Poh Ming-Zher, Frazier Cherise, Kovitch Thropp Eliana, Mynatt Elizabeth D., Bidwell Jonathan, Mai Roberto, LaFrance W. Curt, Blum Andrew S., Friedman Daniel, Loddenkemper Tobias, Mohammadpour-Touserkani Fatemeh, Reinsberger Claus, Tognetti Simone, and Picard Rosalind W. Multicenter clinical assessment of improved wearable multimodal convulsive seizure detectors. *Epilepsia* 58, 11 (????), 1870–1879. DOI: http://dx.doi.org/10.1111/epi.13899

[9] F. FÃijrbass, S. Kampusch, E. Kaniusas, J. Koren, S. Pirker, R. HopfengÃď rtner, H. Stefan, T. Kluge, and C. Baumgartner. 2017. Automatic multimodal detection for long-term seizure documentation in epilepsy. *Clinical Neurophysiology* 128, 8 (2017), 1466 – 1472. DOI: http://dx.doi.org/https://doi.org/10.1016/j.clinph.2017.05.013

[10] Ellen Greenberger, Chuansheng Chen, Julia Dmitrieva, and Susan P. Farruggia. 2003. Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: do they matter? *Personality and Individual Differences* 35, 6 (2003), 1241 – 1254. DOI: http://dx.doi.org/https://doi.org/10.1016/S0191-8869(02)00331-8

[11] B. E. Heldberg, T. Kautz, H. Leutheuser, R. HopfengÃď rtner, B. S. Kasper, and B. M. Eskofier. 2015. Using wearable sensors for semiology-independent seizure detection - towards ambulatory monitoring of epilepsy. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 5593–5596. DOI: http://dx.doi.org/10.1109/EMBC.2015.7319660

[12] Sidney Kennedy. 2016. A Collaborative Investigation of Predictors of Relapse in Major Depressive Disorder: CAN-BIND-1 Extension Study. https://clinicaltrials.gov/ct2/show/NCT02934334. (2016).

[13] Kurt Kroenke, Tara W. Strine, Robert L. Spitzer, Janet B.W. Williams, Joyce T. Berry, and Ali H. Mokdad. 2009. The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders* 114, 1 (2009), 163 – 173. DOI: http://dx.doi.org/https://doi.org/10.1016/j.jad.2008.06.026

[14] Roger S McIntyre, Michael W Best, Christopher R Bowie, Nicole E Carmona, Danielle S Cha, Yena Lee, Mehala Subramaniapillai, Rodrigo B Mansur, Harry Barry, Bernhard T Baune, and others. 2017. The THINC-Integrated Tool (THINC-it) Screening Assessment for Cognitive Dysfunction: Validation in Patients With Major Depressive Disorder. (2017).

[15] M. MiloÅąeviÄĞ, A. Van de Vel, B. Bonroy, B. Ceulemans, L. Lagae, B. Vanrumste, and S. V. Huffel. 2016. Automated Detection of Tonic-Clonic Seizures Using 3-D Accelerometry and Surface Electromyography in Pediatric Patients. *IEEE Journal of Biomedical and Health Informatics* 20, 5 (Sept 2016), 1333–1341. DOI: http://dx.doi.org/10.1109/JBHI.2015.2462079

[16] Poh Ming-Zher, Loddenkemper Tobias, Reinsberger Claus, Swenson Nicholas C., Goyal Shubhi, Sabtala Mangwe C., Madsen Joseph R., and Picard Rosalind W. Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor. *Epilepsia* 53, 5 (????), e93–e97. DOI: http://dx.doi.org/10.1111/j.1528-1167.2012.03444.x

[17] John A. Naslund, Lisa A. Marsch, Gregory J. McHugo, and Stephen J. Bartels. 2015. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. *Journal of Mental Health* 24, 5 (2015), 321–332. DOI: http://dx.doi.org/10.3109/09638237.2015.1019054 PMID: 26017625.

[18] Janssen Research and Development. 2018. A Prospective, Longitudinal, Observational Study to Evaluate Potential Predictors of Relapse in Subjects With Major Depressive Disorder Who Have Responded to Antidepressant Treatment. https://clinicaltrials.gov/ct2/show/NCT02489305. (2018).

[19] Shivkumar Sabesan and Raman Sankar. 2015. Improving long-term management of epilepsy using a wearable multimodal seizure detection system. *Epilepsy & Behavior* 46 (2015), 56–57.

[20] James R. Williamson, Thomas F. Quatieri, Brian S. Helfer, Gregory Ciccarelli, and Daryush D. Mehta. 2014. Vocal and Facial Biomarkers of Depression Based on Motor Incoordination and Timing. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge (AVEC '14)*. ACM, New York, NY, USA, 65–72. DOI: http://dx.doi.org/10.1145/2661806.2661809