Challenges in data capturing and collection for physiological detection of dementia-related difficulties and proposed solutions.

Abstract. Dementia is a neurodegenerative disease which leads to the individual experiencing difficulties in their daily lives. Often these difficulties cause a large amount of stress, frustration and upset in the individual, however identifying when the difficulties are occurring or beginning can be difficult for caregivers, until the difficulty has caused problematic behavior or undeniable difficulty to the person with dementia. Therefore, a system for identifying the onset of dementia-related difficulties would be helpful in the management of dementia. Previous work highlighted wearable computing-based systems for analyzing physiological data as particularly promising. In this paper, we outline the methodology used to perform a systematic search for a relevant dataset. However, no such dataset was found. As such, a methodology for collecting such a dataset and making it publicly available is proposed, as well as for using it to train classification models that can predict difficulties from the physiological data. Several solutions to overcome the lack of available data are identified and discussed: data collection experiments to collect novel datasets; anonymization and pseudonymization to remove all identifiable data from the dataset; and synthetic data generation to produce a larger, anonymous training dataset. In conclusion, a combination of all the identified methods should ideally be employed in future solutions. Future work should focus on the conductance of the proposed experiment and the sharing of the collected data in the manner proposed, with data ideally being collected from as many people as possible with as many different types of dementia as possible.

Keywords: Dementia, data, wearables, physiological, difficulties, BPSD.

1 Introduction

Dementia is an umbrella term to describe several neurodegenerative diseases which cause the person with the disease to experience cognitive impairment and decline [1]. This impairment and decline then causes the person with dementia (PwD) to experience difficulties in their daily lives, such as misplacing items, forgetting appointments, temporal and spatial disorientation, and more [2]. Thus, a PwD will generally progressively become less independent and require more support and aid in living safely and meeting their needs. With the global number of people with dementia currently at 50 million and predicted to reach 130 million by 2050 [3], it is vital that effective methods to manage the disease and the resulting difficulties are found, so that more people with the disease can have greater independence and maintained quality of life for as long as possible [4].

Technology has been shown to be helpful in this regard, and over recent years much work has been done to develop support systems which can aid in the manage-

ment of dementia and the related difficulties [5-11]. In previous work [12], we reviewed the use of wearable computing-based systems for identifying the occurrence of dementia-related difficulties from physiological data. Such systems were found to be able to provide accurate predictions and identifications of difficulties experienced by people with dementia from analysis of physiological data using machine learning. However, we found that no system exists that can predict and aid the management of a comprehensive range of dementia-related difficulties and current systems only predict or provide support for a small number of difficulties in a certain number of scenarios [12]. As such, further work is required to develop a system that can predict and identify a wide range of difficulties experienced by people with various types and stages of dementia. However, to create such a machine learning-based system data is vital, and no available physiological dataset could be found, containing the relevant physiological or behavioral data collected using wearable sensors, that could be used in the project to identify difficulties experienced by people with dementia.

This paper describes the methodology and results of a search for physiological data from people with dementia that was conducted, between June 2019 and December 2020. We also provide a methodology for collecting a physiological dataset from people experiencing dementia-related difficulties, which can be shared with other researchers while protecting the confidentiality of the participants, which would be a novelty as no physiological dataset is currently available in this domain. The rest of the paper is structured as follows. Section 2 provides an overview of the difficulties that a person with dementia may experience followed by a review of the methods used in previous literature to identify those difficulties from physiological data. Section 3 describes the search methodology used in the searches for physiological data that were conducted as part of the project, with the conclusion being that there are no publicly available datasets that meet the requirements of the project. Section 4 proposes a methodology for collecting a physiological dataset from people with dementia and making it available to other researchers. Finally, section 5 provides a conclusion and overview of areas for future work.

2   Background

Much work has been conducted in recent years to develop systems which can identify dementia-related difficulties in a timely and accurate manner. Systems based on wearable computing can be used to collect physiological data to this end, in a passive, non-obstructive manner that is comfortable and convenient for the PwD. One such system for identifying dementia-related difficulties is the BESI system, in which a wrist-worn accelerometer, the Pebble smartwatch, is utilized to track movements of the subjects to detect agitated behaviors [5-7]. In the BESI study, the participant was asked to wear the smartwatch for 30 days, with subject and caregiver dyad numbers ranging from 3 to 10 in each paper and study iteration. Machine learning was then used on the data collected by the wearable to predict the occurrences of agitation. The researchers trained a number of models, such as support vector machines (SVM), adaptive boosting (AdaBoost), and an ensemble of decision trees by bagging (Tree-

Bagging), with the latter providing the most robust prediction of agitated behaviors from unknown data [7]. Though the available literature from the study does not specify the stage of dementia experienced by the participants, it is likely from the community-based setting in which they were tracked that they still lived somewhat independently, or at least in their own home, implying mild to moderate dementia. A wrist-worn device was also used by Melander et al. to collect data from people with dementia in institutionalized settings [11]. The researchers asked several participants to wear an Empatica E4, which tracked the electrodermal activity (EDA) of the participant, while a nurse was asked to record observations of dementia-related agitation and difficulties on a provided chart. The EDA data was then labelled using the observations and the researchers found a high correlation not only with the data recorded at the time the difficulty was observed but also 1 to 2 hours prior to the observation. Thus, EDA data could be used to predict the occurrence of a dementia-related difficulty up to 2 hours before it is observed.

Sefcik et al. focused their research on the later stages of dementia, using a chest-worn ECG sensor to monitor the heart rate of people with advanced dementia who exhibit persistent vocalizations (PV) [13]. PV are described as uncontrolled or disruptive vocalizations with no specific communicative purpose. The participants are asked to wear the device for 2 hours at a time, while caregivers recorded instances of PV. The participants heart rates where then compared on days that they exhibited PV and those they did not, and it was found that heart rate had a strong correlation with PV exhibition, supporting the use of physiological data from wearable computing devices to predict dementia-related difficulties in the moderate to severe stages of dementia. Heart rate was also tracked by Nesbitt et al. in their study to identify and predict dementia-related agitation, in which they used an android smartwatch to track limb movement and heart rate with an accelerometer and PPG sensor, respectively [14]. They also utilized an android smartphone's microphone, worn in a pouch around the neck of the participant. The data collected by the smartwatch was found to correlate with many of the observed instance of agitation, indicating a high degree of accuracy, whereas the data collected from the microphone does not correlate with the agitation, and the researchers posit this is due to background noise making the recordings too noisy to be valuable.

In all of the above cases, regardless of the severity of the dementia or the setting in which the data was collected, the analysis of physiological data was vital to the identification and prediction of dementia-related difficulties. Indeed, a useful, relevant, and complete dataset is the most important and fundamental building block or element of any research in computing related fields of study [15-17]. As such, it is vital that any researchers aiming to develop a machine learning model or method for identifying or predicting the occurrences of dementia-related difficulties, have access to a physiological dataset collected from people with dementia who are experiencing dementia-related difficulties.

3   Dataset search methodology

A search for physiological datasets from people with dementia was conducted between June 2019 and December 2020. The online resources and repositories searched include: UK.gov; data.europa.eu; GitHub Awesome Data; NHS Digital; European Health Information Gateway; reddit.com/r/datasets/; apps.who.int/gho/data/; UK Data Service; Google Search; alzpossible.org; CDC Data sets; Global Open Data Index; LJMU Open Data; biogps.org; niagads.org; nimhgenetics; ondri.ca; Ontario Brain Institute (OBI); alzheimersresearchuk.org. The predefined search terms and keywords include: Dementia, Alzheimer's, movement, activity, action, daily life, instrumental, basic, playing, games, dancing, wearable, sensor, BPSD, smart device, watch, heart rate, actigraph, GSR, EDA, galvanic skin response, electrodermal activity, cognitive impairment, and MCI. The criteria that a dataset had to meet to be useful to the project were: the dataset must contain physiological data, specifically at least one of heart rate, EDA, or limb movement; the dataset must be collected from people with dementia; the datasets must be legally and ethically available and useable for analysis and use in the project.

 Another method employed was contacting authors of papers published on similar projects or in similar domains where physiological data was collected from people with dementia [6, 11, 14, 18, 19]. An email template was designed by the researchers on this project, with the structure being as follows. The first paragraph introduced the researchers, research institution and the current project. The second paragraph discussed the papers or other works that had been read by the researchers to make us aware of the research of the authors being contacted. Naturally, this second section was changed on each occasion and email was sent, with only a general outline being set out in the template. Finally, the third section of the template contained a request for access to the datasets which we wished to gain access to. This section was also changed for each paper, with the name of the datasets and the data features and types we wished to access being different for each of the authors contacted.

   The search began in June 2019 and was continued in the following months until December 2020, with at least one of the repositories being searched every month, and four large-scale searches of all the repositories being undertaken in June 2019, December 2019 - January 2020, and March 2020-April 2020 and November-December 2020. Each large-scale search lasted around a month, giving researchers enough time to fully vet and evaluate every returned dataset for usefulness and applicability to the project's requirements. In January 2020, following the second large-scale search, authors were contacted using the strategy outlined in the previous paragraph to identify and access relevant datasets from researchers in a similar domain to the project, as it had become clear there was likely to be no available relevant dataset on the listed repositories. No more emails were sent to authors after May 2020, as it was clear that funding and/or privacy concerns were a common barrier for all researchers to sharing their data. Furthermore, and more importantly, a new plan of work had to be devised for the project in reaction to the Covid-19 pandemic, which delayed work on the data search. The search ended in December 2020, as it was clear no relevant datasets were

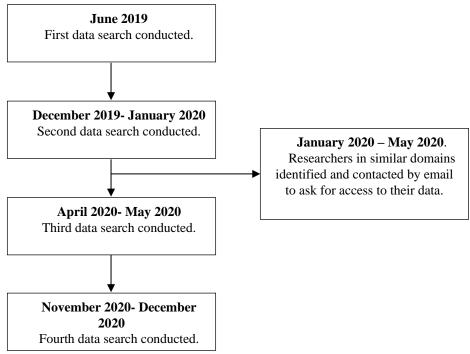available from the repositories or other researchers, and other avenues of work needed to be addressed.

Figure 1. Workflow diagram of the data search methodology.

4 Data search results

No datasets that met the requirement criteria of the project were found during the dataset search outlined in the methodology. The reasons for this are most likely related to the vulnerability of the individuals with dementia from whom the data was collected. People with dementia are classed as vulnerable adults due to their mental impairment [20], which can often make it more difficult to acquire ethical approval to conduct research which includes them as subjects [21]. Furthermore, physiological data collection from people with dementia using wearables can be difficult as the cognitive impairment can lead to problems in collecting the data, especially in the severe stages but also often in mild to moderate dementia [22]. This could be because the subject removes the device either due to discomfort or because the device is unfamiliar to them [8], or because the subject forgets to put a data collection device on [23, 24]. Moreover, caregivers of people with dementia are highly likely to experience stress and other adverse mental or physical burdens due to their caregiving responsibilities, especially in cases where the caregiver is informal i.e. a spouse, family member, or close friend [25-27]. This can also make data collection from people with dementia more difficult than with healthy or non-cognitively impaired subjects, as extra

care must be taken to ensure that any caregivers do not experience any undue burden on top of the burden already commonly experienced. Also, even where data has been collected, it is often difficult to share that data, as the vulnerable status of people with dementia means that their personal data is subject to even greater legal protections than that of an adult not classified as vulnerable [28].

Furthermore, no dataset that met the requirements was found to be or made to be available to the researchers on this project from other authors and researchers in similar domains. Of the researchers who replied, none were willing to share their datasets. There were 2 reasons expressed for not sharing data with the researchers on this project. One of those reasons was privacy or confidentiality concerns, with some researchers stating that the sharing of data from their study would violate the privacy legislation in their respective country or legal jurisdiction or would violate the terms of the agreements signed with participants. Another reason stated for not sharing data was that the sharing of the data would violate the terms of funding agreements, that had clauses which prevented the sharing of the data.

In conclusion, it was concluded that there were no physiological datasets collected from people with dementia experiencing difficulties that were publicly available for use or analysis in the project.

5 Discussion of solutions & proposed methodology

The search for datasets described in the methodology section yielded no useful datasets, and it was concluded that there was no physiological dataset matching the criteria for the project available to be used. In this section, we propose a methodology for developing a prototype support system, including collecting a dataset that does meet the criteria of the project, and is shareable with other researchers in future, so that future research in this domain can be done more easily and cheaply.

5.1 Data collection experiment

The first step in the methodology is performing a data collection experiment in which people physiological data is collected from people with dementia [8]. As the prototype to be developed is targeted to individuals with mild to moderate dementia, the best recruitment channel is likely to be hospital outpatient services or memory clinics [29, 30], or community support groups tailored to individuals with mild to moderate dementia [5-7]. As such, an agreement will be signed with the local memory clinic, and they will identify potential participants from their patient for us based on our inclusion criteria, and we shall request this co-operation in writing. The clinic shall, with the permission of the potential participant, give the details of the potential participant to a qualified medical doctor, who will act as a gatekeeper and approach the potential participant [31]. The gatekeeper will provide the participants with the Participant Information Sheet and go through the protocol and what the study entails with the potential participant and their caregiver(s) and answer any questions that they may have [32]. The inclusion criteria for eligible participants are: a diagnosis of mild

to moderate dementia; living with a caregiver who can record instances of difficulties; have long-term access to a smartphone. The exclusion criterion is that the patient has other neurological disabilities or has significantly reduced ability to complete instrumental activities of daily living due to other non-dementia conditions. Informed and written consent shall be obtained from both the participant and their caregiver for both of their participations, however if the participant is incapable of consenting, their Next of Kin (NoK) or Power of Attorney (POA) will be asked to provide written and informed consent for them, with verbal assent being sought from the participant [11, 13, 14, 33-37]. All consenting and assenting parties will be informed they can withdraw from the study at any time without being required to give a reason or justification.

In previous work we identified the Empatica E4 as the most suitable device available for the proposed data collection experiment [12]. The devices and other items that are required to conduct the experiment are: Empatica E4; Empatica E4 Charger [8]; Smartphone and an Empatica E4 smartphone app; Observational Recording Sheets [11]; Instruction sheets for the participant and caregiver [38]. This instruction sheet will detail what each person should do, as well as providing contact details to the researchers so that they can ask any questions or issues. The participants will receive all the items required, bar the smartphone and app, by the gatekeeper. The gatekeeper will show the participant how the device is to be put on the dominant wrist and how to remove it and correctly charge it. The caregiver will also be shown how to fill in the observational recording sheets and be told when they will be required to do so. The app will be set-up on the smartphone and the phone, and the smartwatch shall be connected via Bluetooth [39].

The data collection will occur for between 2 and 4 weeks. The participant will be asked to wear the device 24 hours a day, except for when it must be removed for charging. The participant will be asked to wear the device 24 hours a day, except for when it must be removed for charging. While deployed on the participants wrist the device shall collect heart rate, electrodermal activity, movement, and skin temperature data from the participant [8]. The caregiver will be asked to record when the participant experiences difficulties, and to record what they were doing at the time and what the difficulty was. To help ensure that fewer potential difficulties are missed, the caregiver shall be sent an alert when the participants heart rate, electrodermal activity or movement is detected to be abnormal, for example if there is a rapid increase in the participants heart rate or is an abnormal amount of movement during the night-time. The caregiver will also be asked to record the activities that the participant has completed each day. This does not need to be 100% accurate and a rough estimation of times and overview of activities is acceptable. The physiological data shall be transferred to the smartphone and then on to the Empatica secure cloud storage service [8, 39]. The observational data recorded on paper shall be stored in the participants home and collected on each visit by the gatekeeper (which will happen once a week).

5.2 Anonymization & Pseudonymization

The next stage of the proposed methodology will be to protect the participants' confidentiality and privacy through anonymization & pseudonymization of the dataset. Anonymization in this instance refers to methods that make the data in the datasets anonymous, so that there is no way to identify the individual participants from which the data was collected [40]. Pseudonymization is slightly different in that it is the replacement of the participants identity with a pseudonym, for example referring to the participants as participant 1, participant 2, and so on [41]. Both methods attempt to protect the identity, privacy, and confidentiality of the individual participants, which is vital to ensuring that any data can be shared and used safely and in line with privacy policies and legislation.

In the proposed methodology, the data will be pseudonymized, with each participant being allocated a number which will be associated with all their data. The only person who will be given access to personal information will be the gatekeeper, with the other members of the research team- who will have no face-to-face interaction with the participants- needing no such knowledge. As the already agreed gatekeeper is a doctor who already has access to the patient records of the potential participants, participant personal data will not be any less confidential than before the experiment. The data shall also be pseudonymized in any future publications of the data or which discuss the data, with the participants being referred to as participant 1, participant 2, etc.

5.3 Synthetic data

Synthetic datasets are datasets that are not collected from participants or situations in the relevant domain, instead being generated synthetically to resemble data from the relevant domain [44]. Synthetic datasets are produced by training a machine learning model, or using one already trained, to identify the statistical and mathematical properties of a dataset and then create a dataset of randomized data that shares the same statistical properties of the original dataset [45].

The use of synthetic data in the domain of dementia-related difficulty identification from physiological data has two major advantages. Firstly, the synthetic dataset will be entirely free from confidentiality or privacy concerns as the data was not collected form people with dementia and is merely computer generated to resemble such data [46]. Furthermore, collection of data from people with dementia is difficult and time consuming, and in many instances, it may be difficult to acquire the number of recruited participant or data collection devices to collect a large dataset. This could potentially lead to overfitting of the results of data analysis to a few small dataset, therefore low generalizability of those results or trained models. Generation of synthetic data with a few small differences to the original to represent different scenarios may help to prevent this overfitting problem and increase the generalizability of the results and trained models [45]. As such, this method shall also be employed in the proposed methodology to increase the volume of data available to train the required machine learning model.

5.4 Model training & data sharing

Once all the data is collected and labelled, and a synthetic dataset has been generated, it shall be used to train several classification models with the MATLAB software [47]. This will allow for the identification of the best model for identifying and predicting the difficulties from the physiological dataset. This model will then be used in the development of a prototype which can identify the difficulties and provide an automatic digital intervention.

Furthermore, the collected and generated data will be completely anonymous and pseudonymized, and consent will have been sought from the participants for any collected data to undergo secondary analysis, after it has been entirely de-identified. This means that the data can be published on the host institution's secure data repository [48], meaning other researchers in this domain can access the data on request. This would be novel for the domain, as no other dataset of this nature is currently publicly available to researchers in this domain.

6   Conclusion

In conclusion, there is currently a lack of physiological datasets pertaining to dementia-related difficulties or from people with dementia, and this is impeding progress in the field of assistive technology for the management of dementia. Therefore, many people with dementia and their caregivers may miss out on the benefits of such a system due to increased time needed to research and develop them. It is thus vital that a relevant physiological dataset is collected, processed, and made available as soon as possible, so that further research can be done on this domain, and all the benefits of the proposed system can be realized by people with dementia and their caregivers- including increased QoL and independence, and reduced burden for caregivers- within the near future.

Such an experiment should be conducted in a way that provides the highest possible levels of privacy and confidentiality for the participant, while also getting their express permission to share a processed version of their data in a manner that does not compromise their privacy. One solution to ensure this is the use of anonymization and pseudonymization, both of which remove all personal data from the physiological data, and thus makes the resulting dataset non-personal data, thus more ethical to share. Finally, synthetic data generation can be used to further protect participant privacy and confidentiality, and to generate larger training datasets with which to train machine learning models. As such the proposed methodology shall include: a data collection experiment to collect an initial physiological dataset; anonymization or pseudonymization to protect the identity and confidentiality of participants; and synthetic dataset generation to provide a larger training dataset for machine learning. Once the data is labelled and completely de-identified it will be used to train a variety of machine learning models and it will be shared via a secure data repository, becoming a novel, publicly available dataset.

Future work in this area should include the conducting of the proposed data collection experiment, in which physiological data should be collected from people with

dementia. It should be ensured that participants consent to the sharing and secondary analysis of the data collected from them, with that data sharing and secondary analysis being completing in a manner that protects participant confidentiality. Ideally, data should be collected from as many people with dementia, including people with different types and severity of dementia, as is possible, which will allow for greater generalizability of results from data analysis, as well as allowing research to be done on supporting individuals with every type and severity of dementia.

References

[1] L. Koumakis, C. Chatzaki, E. Kazantzaki, E. Maniadi, and M. Tsiknakis, "Dementia Care Frameworks and Assistive Technologies for Their Implementation: A Review," *IEEE reviews in biomedical engineering,* vol. 12, pp. 4-18, 2019.

[2] K. Okabe *et al.*, "Effects of neuropsychiatric symptoms of dementia on reductions in activities of daily living in patients with Alzheimer's disease," *Geriatrics & gerontology international,* vol. 20, no. 6, pp. 584-588, 2020.

[3] WHO. "Dementia." World Health Organisation. https://www.who.int/news-room/fact-sheets/detail/dementia (accessed 5th May, 2020).

[4] R. Howard *et al.*, "The effectiveness and cost-effectiveness of assistive technology and telecare for independent living in dementia: a randomised controlled trial," *Age and Ageing,* 2021.

[5] R. Alam *et al.*, "Motion biomarkers for early detection of dementia-related agitation," in *Proceedings of the 1st Workshop on Digital Biomarkers*, 2017, pp. 15-20.

[6] R. Alam, M. Anderson, A. Bankole, and J. Lach, "Inferring physical agitation in dementia using smartwatch and sequential behavior models," in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 2018: IEEE, pp. 170-173.

[7] R. Alam, A. Bankole, M. Anderson, and J. Lach, "Multiple-Instance Learning for Sparse Behavior Modeling from Wearables: Toward Dementia-Related Agitation Prediction," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019: IEEE, pp. 1330-1333.

[8] F. Amato *et al.*, "CLONE: a Promising System for the Remote Monitoring of Alzheimer's Patients: An Experimentation with a Wearable Device in a Village for Alzheimer's Care," in *Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good*, 2018, pp. 255-260.

[9] B. Kikhia *et al.*, "Utilizing a wristband sensor to measure the stress level for people with dementia," *Sensors,* vol. 16, no. 12, p. 1989, 2016.

[10] B. Kikhia *et al.*, "Utilizing ambient and wearable sensors to monitor sleep and stress for people with BPSD in nursing homes," *Journal of Ambient Intelligence and Humanized Computing,* vol. 9, no. 2, pp. 261-273, 2018.

[11] C. Melander, J. Martinsson, and S. Gustafsson, "Measuring electrodermal activity to improve the identification of agitation in individuals with dementia," *Dementia and geriatric cognitive disorders extra,* vol. 7, no. 3, pp. 430-439, 2017.

[12] M. Harper and F. Ghali, "A Systematic review of wearable devices for tracking physiological indicators of Dementia-related difficulties," presented at the Developments in E-Systems, Online, 2020.

[13] J. S. Sefcik, M. Ersek, J. R. Libonati, S. C. Hartnett, N. A. Hodgson, and P. Z. Cacchione, "Heart rate of nursing home residents with advanced dementia and persistent vocalizations," *Health and Technology,* pp. 1-5, 2019.

[14] C. Nesbitt, A. Gupta, S. Jain, K. Maly, and H. R. Okhravi, "Reliability of wearable sensors to detect agitation in patients with dementia: A pilot study," in *Proceedings of the 2018 10th International Conference on Bioinformatics and Biomedical Technology*, 2018, pp. 73-77.

[15] T. Zhou, Z. Song, and K. Sundmacher, "Big data creates new opportunities for materials research: A review on methods and applications of machine learning for materials design," *Engineering,* vol. 5, no. 6, pp. 1017-1026, 2019.

[16] N. Huck, "Large data sets and machine learning: Applications to statistical arbitrage," *European Journal of Operational Research,* vol. 278, no. 1, pp. 330-342, 2019.

[17] I. Lee and Y. J. Shin, "Machine learning for enterprises: Applications, algorithm selection, and challenges," *Business Horizons,* vol. 63, no. 2, pp. 157-170, 2020.

[18] C. Lai Kwan, Y. Mahdid, R. Motta Ochoa, K. Lee, M. Park, and S. Blain-Moraes, "Wearable Technology for Detecting Significant Moments in Individuals with Dementia," *BioMed research international,* vol. 2019, 2019.

[19] L. Valembois, C. Oasi, S. Pariel, W. Jarzebowski, C. Lafuente-Lafuente, and J. Belmin, "Wrist actigraphy: a simple way to record motor activity in elderly patients with dementia and apathy or aberrant motor behavior," *The journal of nutrition, health & aging,* vol. 19, no. 7, pp. 759-764, 2015.

[20] K. S. Thomas, W. Zhang, P. Y. Cornell, L. Smith, B. Kaskie, and P. C. Carder, "State variability in the prevalence and healthcare utilization of assisted living residents with dementia," *Journal of the American Geriatrics Society,* vol. 68, no. 7, pp. 1504-1511, 2020.

[21] C. Benson, A. Friz, S. Mullen, L. Block, and A. Gilmore-Bykovskyi, "Ethical and Methodological Considerations for Evaluating Participant Views on Alzheimer's and Dementia Research," *Journal of Empirical Research on Human Research Ethics,* p. 1556264620974898, 2020.

[22] M. Kaenampornpan, N. D. Khai, and K. Kawattikul, "Wearable Computing for Dementia Patients," in *International Conference on Computing and Information Technology*, 2020: Springer, pp. 21-30.

[23] E. Grober, D. Wakefield, A. R. Ehrlich, P. Mabie, and R. B. Lipton, "Identifying memory impairment and early dementia in primary care," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,* vol. 6, pp. 188-195, 2017.

[24] L. McGarrigle, S. E. Howlett, H. Wong, J. Stanley, and K. Rockwood, "Characterizing the symptom of misplacing objects in people with dementia: findings from an online tracking tool," *International psychogeriatrics,* vol. 31, no. 11, pp. 1635-1641, 2019.

[25] M. H. Connors, K. Seeher, A. Teixeira-Pinto, M. Woodward, D. Ames, and H. Brodaty, "Dementia and caregiver burden: A three-year longitudinal study," *International journal of geriatric psychiatry,* vol. 35, no. 2, pp. 250-258, 2020.

[26] A. P. Allen *et al.*, "Informal caregiving for dementia patients: the contribution of patient characteristics and behaviours to caregiver burden," *Age and ageing,* vol. 49, no. 1, pp. 52-56, 2020.

[27] J.-A. Su and C.-C. Chang, "Association between family caregiver burden and affiliate stigma in the families of people with dementia," *International journal of environmental research and public health,* vol. 17, no. 8, p. 2772, 2020.

[28] B. S. Husebo, H. L. Heintz, L. I. Berge, P. Owoyemi, A. T. Rahman, and I. V. Vahia, "Sensing technology to facilitate behavioral and psychological symptoms and to monitor treatment response in people with dementia. A Systematic Review," *Frontiers in pharmacology,* vol. 10, p. 1699, 2020.

[29] A. Karakostas, I. Lazarou, G. Meditskos, T. G. Stavropoulos, I. Kompatsiaris, and M. Tsolaki, "Sensor-based in-home monitoring of people with dementia using remote web technologies," in *2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*, 2015: IEEE, pp. 353-357.

[30] NHS. "How to get a dementia diagnosis." NHS.uk. https://www.nhs.uk/conditions/dementia/diagnosis/ (accessed April 19th, 2020).

[31] R. Bartlett, R. Milne, and R. Croucher, "Strategies to improve recruitment of people with dementia to research studies," *Dementia,* vol. 18, no. 7-8, pp. 2494-2504, 2019.

[32] J. Waite, F. Poland, and G. Charlesworth, "Facilitators and barriers to co-research by people with dementia and academic researchers: Findings from a qualitative study," *Health Expectations,* vol. 22, no. 4, pp. 761-771, 2019.

[33] S. S. Khan *et al.*, "Agitation Detection in People Living with Dementia using Multimodal Sensors," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019: IEEE, pp. 3588-3591.

[34] S. Spasojevic *et al.*, "A pilot study to detect agitation in people living with dementia using multi-modal sensors."

[35] D. Goerss *et al.*, "Automated sensor-based detection of challenging behaviors in advanced stages of dementia in nursing homes," *Alzheimer's & Dementia,* 2019.

[36] S. Teipel *et al.*, "Multidimensional assessment of challenging behaviors in advanced stages of dementia in nursing homes—The insideDEM framework," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,* vol. 8, pp. 36-44, 2017.

[37] N. Vuong, S. Chan, C. T. Lau, S. Chan, P. L. K. Yap, and A. Chen, "Preliminary results of using inertial sensors to detect dementia-related wandering patterns," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015: IEEE, pp. 3703-3706.

[38] A. Page, K. Potter, R. Clifford, A. McLachlan, and C. Etherton-Beer, "Prescribing for Australians living with dementia: study protocol using the Delphi technique," *BMJ open,* vol. 5, no. 8, p. e008048, 2015.

[39] Empatica. "E4 Wristband." Empatica. https://www.empatica.com/research/e4/?utm_source=Google&utm_medium=cpc&utm_ca mpaign=conversion&gclid=Cj0KCQiAvc_xBRCYARIsAC5QT9l9EUAmaWXCRxEQ17 aUYewwgoMkYIn-xFsuJld5R-Ib_6wGuhYqKToaAsomEALw_wcB (accessed 30th January, 2020).

[40] S. Murthy, A. A. Bakar, F. A. Rahim, and R. Ramli, "A comparative study of data anonymization techniques," in *2019 IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, 2019: IEEE, pp. 306-309.

[41] L. Bolognini and C. Bistolfi, "Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation," *Computer Law & Security Review,* vol. 33, no. 2, pp. 171-181, 2017.

[42] A. Deshpande, "Sypse: Privacy-first Data Management through Pseudonymization and Partitioning."

[43] J. Menner, P. Lewandowska, and F. Zabel, "THE IMPACT OF DATA PRIVACY REGULATIONS ON DRUG UTILIZATION DATA SHARING FOR INNOVATIVE PRICING ARRANGEMENTS," in *Value in Health*, 2018, vol. 21: ELSEVIER SCIENCE INC STE 800, 230 PARK AVE, NEW YORK, NY 10169 USA, pp. S211-S211.

[44] K. Nikolaidis, S. Kristiansen, V. Goebel, T. Plagemann, K. Liestøl, and M. Kankanhalli, "Augmenting physiological time series data: A case study for sleep apnea detection," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 2019: Springer, pp. 376-399.

[45] A. Yale, S. Dash, R. Dutta, I. Guyon, A. Pavao, and K. P. Bennett, "Generation and evaluation of privacy preserving synthetic health data," *Neurocomputing,* vol. 416, pp. 244-255, 2020.

[46] H. Ping, J. Stoyanovich, and B. Howe, "Datasynthesizer: Privacy-preserving synthetic datasets," in *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, 2017, pp. 1-5.

[47] A. L. Manuel *et al.*, "Interactions between decision-making and emotion in behavioral-variant frontotemporal dementia and Alzheimer's disease," *Social cognitive and affective neuroscience,* vol. 15, no. 6, pp. 681-694, 2020.

[48] LJMU. "Welcome to LJMU Data Repository." LJMU. http://opendata.ljmu.ac.uk/ (accessed 31st March, 2021).