Thesis Portfolio

Behavioral and Environmental Sensing and Intervention for Cancer
(Technical Topic)

Effects of Predictive Algorithms in Opioid Pain Management
(STS Topic)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Penn Bauman
Spring, 2022

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

The experience of advanced cancer comes with many struggles, but one of the most constant is the task of pain management. Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) is a join research project between the University of Virginia School of Nursing and School of Engineering and Applied Science, which aims to use predictive algorithms to simplify the process of managing cancer pain at home. The final goal of the project is to create a system which can monitor cancer patients day-to-day activities with noninvasive sensors and use that data to predict pain events. Predictions will be made by a machine learning algorithm which can analyze patients data in real time. With these predictions, the system can automate the timing of pain medication, reducing the burden of pain management for patients and their caregivers.

In order to create this predictive algorithm and allow patients to use it, monitoring systems must be created to collect data about patients and their environments. A prototype of this system is currently being developed and tested by the BESI-C project. This system consists of cloud infrastructure to store and analyze data as well as devices which are placed in patient homes, including wall-mounted sensor packs, smart watches, a devices to allow the system internet access. This system allows data to be collected from patients and their homes continually and automatically. The technical portion of this thesis describes in detail how the wall-mounted sensors work and the other technologies which support them work.

If completed and put into use outside of academic studies, the BESI-C system would be unlike and existing medical product or device. This makes it difficult to predict possible social side effects or ethical issues that might appear during its use. To predict, and hopefully prevent, some of these issues before they appear, I have considered ethical issues generated by research into systems similar to the BESI-C system. To consider the effects of BESI-C system's constant monitoring, I researched self tracking systems. To compare BESI-C system to existing solutions to support patients dealing with advanced cancer, I consider research into ethical issues surrounding inpatient care. Finally, to predict ethical issues that might arise from the BESI-C project's use of machine learning, I considered research into algorithm bias. With the ethical concerns generated by these different fields of research, I am able to predict multiple possible ethical issues which could arise from the BESI-C system. The ethical concerns generated by these different fields of research predict multiple possible ethical issues which could arise from the BESI-C system. Knowing potential issues in advance, I can recommend strategies the BESI-C project should focus on to minimize or avoid their impacts, creating a more ethical and more helpful product.

Behavioral and Environmental Sensing and Intervention for Cancer Sensor System

CS4991 Capstone Report, 2021

Penn Bauman
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia, USA
pcb8gb@virginia.edu

ABSTRACT

Opioid painkillers are an essential part of modern pain management, especially for patients with serious illnesses such as cancer, but their use carries with it the risk of serious addiction. Our project aims to use machine learning to reduce the risk of addiction with opioid painkillers, without limiting their effectiveness or causing patients to experience unnecessary pain. To ensure this model can gain the information required to properly assess the each patient, data must be collected both about the patient and their environment. As part of this data collection I worked on a system of sensor packs placed in the patient's home to collected environmental data. Each of the sensor packs is configured with a set of scripts and then runs a program which sends collected data to a central server, where it can be analyzed. The prototypes of these sensor packs produced are currently being tested in real patient homes, and are performing their tasks successfully.

1 INTRODUCTION

For years opioid addiction has ravaged many communities throughout the United States, killing hundreds of thousands of people [1]. As this epidemic has become more apparent, numerous improvements have been made to our medical system to decrease the chances of opioid painkillers leading to addiction, but there are still many for whom it is a serious risk. One group of people with such a risk is cancer patients using opioids to treat chronic pain, a relatively common situation due to no alternative high-strength painkillers. In these long-term pain management situations, avoiding addiction requires carefully managing the frequency and dosing of medication to ensure the patient is comfortable without unnecessary drug intake. This balance can rarely be maintained perfectly, as variation in various biological systems each data causes inevitable difficulties. This leaves doctors and patients stuck between giving patients enough medication to prevent any pain, which can result in excess medication, or ensuring the patient is never over-medicated, making breakthrough pain inevitable. The project this research is based on attempts to reduce both breakthrough pain and over-medication by using machine learning to dose medication based on real-time patient data.

Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) is a joint research project of the University of Virginia School of Engineering and Applied Science and Nursing School [2]. The project is funded by the National Institutes of Health and is currently under active research and development, with initial versions of the data collection system being tested with real patients. The system aggregates environmental, biometric, and qualitative experiential data which is then utilized by a personalized machine learning algorithm to predict when pain events will appear and when medication is required to prevent them. If successful, this predictive system could allow effective pain suppression with fewer total opioids consumed than traditional dosing methods, reducing the chances of addiction.

2 BACKGROUND

To generate the pain prediction to be used by the BESI-C system a number of component systems must be developed. These include patient and environmental data collection systems, data analysis systems, and ultimately a machine learning model utilizing this data to make the actual predictions. I have focused my work for the project on the environmental data collection system, which consists of sensor pack and the network infrastructure to transmit their data. These elements are designed to take a selection of data readings at regular intervals continuously [3]. This means that the devices and software developed need to be capable of reading data consistently for long periods of time, and must be able to transmit that data reliably.

3 PROJECT DESIGN

The environmental data collection system for BESI-C consists of 3 major parts: the sensor packs which actually collect data, a central base station which connects all devices in the patient home, and the remote server these devices connect to [4]. Each of these components interact with the other two in various ways, so their development much be coordinated and synchronized. For the servers interaction with the sensor packs and base stations this interaction is standardized through the remote servers API, which is the same for all devices. The interaction between base stations and sensor packs is done through the base stations status as a WiFi router, so it can be standardized using the WiFi standard.

3.1 Sensor Packs

Environmental data for the BESI-C system is collected in each patient home through four sensor packs each placed in a different room. Each sensor pack collects four simple data points each second: the temperature, the barometric pressure, the humidity, and the light intensity [4]. These simple data points are sent to the remote server every time they are collected allow researchers live access to the data being produced. Additionally, each one collects half second samples of sound and processes them locally to extract 32 audio features which can be used from analysis without recording patients and violating their privacy. All of this data is collected in ZIP files on each sensor pack, these ZIP files are uploaded to AWS S3 once every 30 minutes.

Each of these packs is built using a Raspberry Pi Zero W, wired to a custom printed sensor board combining various sensors. The Raspberry Pi runs the Raspberry Pi OS Linux distribution, configured to interface with the sensors. This configuration install the necessarily drivers and sets the necessarily configuration options to allow data collection, and also sets up a system of scripts which keep the sensor pack in regular communication with the remote server. This communication allows the sensor packs to transmit data to the remote server, but also allows the remote server to change if a sensor pack is assigned to a patient home or not. Each sensor pack checks its assignment every 30 minutes, and if it is assigned to a patient home receives a deployment name and relay id. There pieces of data are included with all subsequent data uploads to AWS S3 allow data analysis to know where each piece of data came from.

3.2 Base stations

The each patient home is also given a base station which primarily provides network access to other devices in the home. These base stations are Raspberry Pi 4Bs which run Raspberry Pi OS, and are configured as WiFi routers. Base stations are then connected to a mobile hotspot, allowing sensors packs which connect to the base station's WiFi network to access the remote server. Base stations also regular communicated directly with the remote server, so that the connection state of each base station can be monitored.

3.3 Remote Server

The remote server controls and coordinates all base stations and sensor packs used by BESI-C. It maintains a list of all currently active devices, along with information about the last data the collected, last time they were connected, the last IP address they connected from, and if they are currently assigned to a patient home. Some of this data is produced by the devices and then sent the server where it is stored for research use, while information about which devices are assigned to patients is set on the server and then communicated to devices. All of these interactions are done through a REST API.

4 RESULTS

This system is still being tested, but it is currently preforming well with consistent data transmission most of the time. The system does continue to have occasional issues with temporary connection failures. These issues do not result in data lost, because data is still collected in the ZIP files, which can be uploaded after the connection has resumed.

5 FUTURE WORK

The BESI-C project is still an active research project at this time, and we will continue to iterate on and improve the current version of the system. A major part of this future work will involve updating the system to work with new sensor hardware, a change required by the current global shortage of electric components. This change will require that the components of the system which read and process data on each of the sensor packs be completely re-engineered. These new versions of the sensor packs will continue be compatible with the current version of the remote monitoring system, allowing both current versions and future versions of the sensor packs to be used at the same time.

We may also decide to change how the relays function, by adding functionality such as adding new sensors or more radical changes such as making them battery powered. If new sensors are added, new code with have to be written to read and manage them, but the overall structure of the system will not need to be changed. If sensor packs are changed to use battery power more significant modifications will be required. Currently the audio analysis required to maintain patient anonymity is the most significant power draw in the sensor packs, so this will likely have to be eliminated to allow for battery power operations. This could be done by no longer collecting audio data, if it is found not to be useful, or by streaming raw audio data to the base station and processing it there.

REFERENCES

[1] M. R. Jones, O. Viswanath, J. Peck, A. D. Kaye, J. S. Gill, and T. T. Simopoulos, "A brief history of the opioid epidemic and strategies for pain medicine," *Pain and Therapy*, vol. 7, no. 1, pp. 13–21, Jun. 2018, ISSN: 2193-8237, 2193-651X. DOI: 10.1007/s40122-018-0097-6. [Online]. Available: http://link.springer.com/10.1007/s40122-018-0097-6 (visited on 10/18/2021).

[2] V. Lebaron, *BESI-C*, 2021. [Online]. Available: https://besic.org.

[3] V. LeBaron, R. Bennett, R. Alam, *et al.*, "Understanding the experience of cancer pain from the perspective of patients and family caregivers to inform design of an in-home smart health system: Multimethod approach," *JMIR Formative Research*, vol. 4, no. 8, e20836, Aug. 26, 2020, Company: JMIR Formative Research Distributor: JMIR Formative Research Institution: JMIR Formative Research Label: JMIR Formative Research Publisher: JMIR Publications Inc., Toronto, Canada. DOI: 10.2196/20836. [Online]. Available: https://formative.jmir.org/2020/8/e20836 (visited on 11/17/2021).

[4] V. LeBaron, J. Hayes, K. Gordon, *et al.*, "Leveraging smart health technology to empower patients and family caregivers in managing cancer pain: Protocol for a feasibility study," *JMIR Research Protocols*, vol. 8, no. 12, e16178, Dec. 9, 2019, Company: JMIR Research Protocols Distributor: JMIR Research Protocols Institution: JMIR Research Protocols Label: JMIR Research Protocols Publisher: JMIR Publications Inc., Toronto, Canada. DOI: 10.2196/16178. [Online]. Available: https://www.researchprotocols.org/2019/12/e16178 (visited on 10/17/2021).

Social Costs of Predictive Algorithms in Opioid Pain Management

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Penn Bauman
Spring, 2022

On my honor as a niversity Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature ______________________________ Date __________
Penn Bauman

Signature ______________________________ Date __________
Joshua Earle, Department of Engineering and Society

Introduction

The experience of advanced cancer comes with many struggles, but one of the most constant is the task of pain management. In many cases, patients end up taking multiple pain medications and the organization of these different prescriptions can become another difficulty in an already trying experience. Serious chronic cancer pain is commonly treated with two types of opioid pain medication at the same time: daily slow release opioids which provide baseline pain management and quick release opioids which can be used to treat more sudden, so called ”breakthrough pain” (Silvia Gonella et al., 2019). It can be difficult for patients, and for the informal family caregivers who assist them, to time these quick release pain killers well. The reactive nature of these interventions means that patients commonly suffer pain for a period before the medication has time to take effect. While inpatient care or the hiring of professional care giving staff can alleviate the burden of medication management for patients, these options are not always financially feasible or how a patient wants to spend their time. Inpatient care in particular can have significant downsides for patients, causing many cancer patients to choose to stay at home in an environment where they are comfortable. This leaves many patients and family caregivers dealing with the job of managing opioid pain medication on top of the many other struggles of advanced illness (Chi & Demiris, 2017).

Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) is a join research project between the University of Virginia School of Nursing and School of Engineering and Applied Science, which aims to use predictive algorithms to simplify the process of managing cancer pain at home. The final goal of the project is to create a system which can monitor cancer patients day-to day activities with noninvasive sensors and use that data to predict breakthrough pain events. Predictions will be made by a machine learning algorithm which can analyze patients data in real time. With these predictions, a system can be designed to automate the timing of pain medication, reducing the burden of pain management for patients and their caregivers.

The BESI-C project has already produced a prototype monitoring system to begin the process of collecting training data. This system consists of cloud infrastructure to store and analyze data as well as equipment boxes which are shipped to patients participating in the research study. Each box contains 4 wall-mounted sensor packs; 2 pairs of smart watches, worn by the patients and caregivers, respectively; and a mobile hotspot, raspberry pi 4 computer, and cellphone which allow the other devices to connect to the internet (LeBaron et al., 2020). These devices allow data to be collected at all times and automatically uploaded to cloud services where it can be processed an analyzed.

The system collects five groups of data: environment sensors data, extracted audio feature data, biometric sensors data, Bluetooth proximity data, and patient questionnaire data. Environmental sensor data and extracted audio feature data are both collected by the wall-mounted sensor packs. The environmental data collected once per second and consists of temperature, barometric pressure, humidity, and light intensity readings. The extracted audio data is designed to allow analysis of noise in patients homes without compromising HIPAA privacy rights by recording and transmitting audio reading. These numerous features, including average audio frequency and audio amplitude, are extracted on the wall-mounted sensors without audio data ever being saved to persistent storage. Biometric data is collected by the Android watches from the built-in heart rate monitor, pedometer, and accelerometer. Bluetooth proximity data is collected each second by the Android watches for each of the wall sensors, allowing the watches to be located within the home. Finally, patient questionnaire data is collected through a set of always-available and automatically timed surveys presented to patients through a custom app on the Android watches (LeBaron et al., 2019).

If the BESI-C system is completed and put into use outside of academic studies, it will be unlike and existing medical product or device. No medical system yet uses machine learning to manage something as essential as opioid pain medication. To predict and consider possible societal consequences and ethical issues which could arise from such a unique system, I will consider existing research on similar systems. These related systems are each similar to a different aspect of the BESI-C system, so that the greatest breadth of possible ethical issues can be considered. In order to gain insight into possible side effects of the BESI-C system's constant monitoring of patients, I will consider research into the growing category of self-monitoring systems. It is also important to compare the BESI-C system to other systems which assist cancer patients with managing their care and pain. For this I will consider the ethical concerns and complexities created by inpatient cancer care. Finally, the use of machine learning in the BESI-C system must be analyzed for ethical issues. To predict issues with BESI-C's algorithm, I will consider other research on machine learning and particularly on its use in the medical field. By considering ethical issues raised in research about each of these socio-technical systems, I can predict a variety of potential ethical issues created by the BESI-C system. If issues can be predicted before they appear, I can consider ways to mitigate them and potentially avoid some of them before they even occur.

Analysis

Quantified Self

The first related system I will consider for ethical issues are self-tracking systems, where individuals track their own lives and habits. This self-tracking, referred to collectively in research as the qualified self, can be as simple as a handwritten log of events or as complex as aggregating multiple sets of data from smartphone sensors or home medical equipment (Swan, 2013, p. 1). This data can be used to track health over time or to correlate behaviors with particular outcomes, allowing harmful behaviors to be avoided or certain aspects of ones life optimized. While not self-operated, the BESI-C system is quite similar to many Quantified Self systems, making Quantified Self research valuable in anticipating possible social outcomes of the system. The central ethical issues which appear in research about the Quantified Self are those of user privacy and autonomy (Sharon, 2017, p. 106).

Privacy in particular has become a central issue when digital self-tracking systems handle large quantities of user data, which has become increasingly common. Services from large companies such as the iOS health app as well as smaller services which offer more targeted self-tracking abilities, both collect and store user data in huge aggregated databases. These databases are always a threat to user privacy, but these issues become much more pressing when such data is revealed to the public intentionally through ”social media” like features, even if such data is anonymized. In one notable incident, complete anonymous data made available by a run tracking was able to reveal the locations of secret US Army bases in Syria (Hern, 2018).

While both Quantified Self systems and the BESI-C system collect regular data about people’s daily lives, they are also different in important ways. The official medical status of the BESI-C project means that data collected by the system is subject to HIPAA regulations which require strict data privacy. This means that BESI-C maintains much stricter user privacy than most apps used for self-monitoring (Yaraghi & Gopal, 2018).

The other major area of ethical concern with Quantified Self systems is over users potential lose of autonomy. Many systems considered examples of the Quantified Self use simple formulas to convert their collected data into directions for users. If incorrect these directions could push people to make unhealthy decisions or control their lives unnecessarily. BESI-C certainly similar ethical concerns by allowing its predictive algorithms to govern the timing of patient’s medication. While users will retain final control over their actual consumption of medication, they are likely to respond to the BESI-C systems manipulation of this process in a variety of ways. Patients are likely to respond to these new sources of control in varying ways. Some users are likely to embrace such a system, particularly if the system works well for them. A few patients may even actively involve themselves in tuning and evaluating the system, as occurs with some pacemaker recipients (Oudshoorn, 2015, p. 69). However there will be other patients who will be uncomfortable with a computer making health choices for them, especially older patients who are less comfortable with technology. As the project continues these types of responses must be taken into account, and if possible the system should be tweaked to prevent users from experiencing the system as too overbearing.

Inpatient Care

(Ward et al., 1993) While the BESI-C system is designed for home care based management, it is valuable to consider ethical issues brought up by inpatient cancer care. BESI-C plans to significantly increase the presence of medical equipment in home making it useful to consider the inpatient care which also situates patients in an environment filled with medical equipment. Inpatient palliative care also relieves the burden of managing pain medication from cancer patients and their caregivers, even more completely than the BESI-C system does. Like inpatient care professional, the BESI-C system may be perceived as wielding some level of medical expertise and the effects this could have on patients both for good or bad should be considered. The presence of such perceived expertise may comfort some patients, relieving anxiety they may feel about the complex task of deal with cancer, but this perceived expertise could also limit patients willingness to advocate for themselves (Ward et al., 1993). If they expect the BESI-C system to be able to fully manage their pain without their input, they may be inclined to provide limited feedback, potential reducing the effectiveness of the very system they are putting their trust in.

One major ethical complexity with inpatient pain management is its restriction of patient privacy. This lack of privacy is largely due to the nature of inpatient care which requires that doctors, nurses, and other hospital staff enter the patients room regularly, necessary invading privacy each time (Street & Love, 2005). This issue becomes especially serious in facilities that do not have enough room to give patients private rooms causing patients to spend all of their time in a shared space without real privacy. The BESI-C system would be able to completely alleviate these issues, allowing patients to enjoy the full privacy of their own homes.

However, the BESI-C system does bring with it the ethical concerns created by constant monitoring from medical equipment. Although more subtle than the presence of medical professionals, medical monitoring equipment can also violate user privacy. The BESI-C system actually intensifies the risk of these ethical issues, because it transfers and store data remotely, making it much more vulnerable. Luckily all of the data collected by both inpatient care equipment and the BESI-C system is HIPAA protected, preventing it from being shared and require that it be store without identifying information.

Another major ethical concern in inpatient cancer care is the loss of patient autonomy, as many of the details of their treatment and life are managed directly by medical professionals. While doctor and nurses are strive to respect patients wishes and consider their opinions in how their care is managed, they may infringe on patient autonomy in may inadvertent ways (Niemeyer-Guimarães & Schramm, 2017). Many of these infringements are the result of hospitals practical constraints, certain procedures must be done in certain ways to appease regulations and resources must be managed and allocated to ensure the entire system functions smoothly, even if this may not be ideal for specific patients. BESI-C also largely avoids this issue with inpatient care by keeping the patient in their own home where they and there family preform these organizational tasks. The ability for patients to control these tasks is not always beneficial however, as these task may be overwhelming for some patients or their families, especially older patients or those with very advanced cancers (Chi & Demiris, 2017). Patients may therefore sometimes be better served having these aspects of their life managed for them, allowing them to better relax.

Machine Learning

It is also important to consider ethical issues created by machine learning systems, which BESI-C plans to utilize. The use of these algorithms has exploded in recent years, allowing computer to efficiently solve numerous problems which were previously too complex to be practical to solve. The BESI-C system hopes to make use of some of these cutting edges techniques to produce an algorithm which can predict breakthrough pain in cancer patients. This type of complex application which will requires the utilization of extremely complex and subtle pattern is a perfect use of machine learning technology, without which the project might not be feasible.

However machine learning algorithms can be dangerous as well as useful, as various kinds of algorithmic bias can occur and prevent the algorithm's proper function. Machine learning is particularly susceptible to bias because systems are largely black boxes which even their creators don't understand. This make is easy for bias to go unnoticed because its causes are concealed.

Algorithmic bias typically develops because algorithms are created and tested on bias data, either because only limited data was used or because data inherits some bias from the real world (O'Neil, 2016). This can make machine learning models useless to more marginalized communities who the system's creators did not consider during its design. To ensure that the BESI-C system is useful to everyone, it is essential that bias is avoided in its algorithm.

With medical machine learning systems it is particularly common for training data to fail to include data from specific populations. For example, if a algorithm designed to identify skin cancer is train using examples of skin cancer on light skin only, it may fail to identify or misidentify skin cancer of darker skin (Vokinger et al., 2021). The BESI-C system could produce this kind algorithm bias because of the geographically limited set of patients the early version of the system are available to. Currently the system is being tested only with patients in the region surrounding Charlottesville, which will limit the diversity of the training data collected. To ensure that this type of bias does not emerge in the final BESI-C system, more diverse training data should collected before the prediction algorithm is created. The system should also be tested on a diverse group of patients and its efficacy monitored not only overall but also for demographic subgroups to ensure equal effectiveness (McCradden et al., 2020).

Conclusion

Privacy was an ethical issue for multiple of the related systems I considered, and it is important to consider if the BESI-C system will create similar issues. The main privacy concern with the BESI-C system is its collection of patient data. This data is all protected by HIPAA regulations, so it must be collected and stored without identifying information, and kept strictly private. While not as much of a privacy concern as some related systems, the BESI-C system still does create risks to user privacy. Any system which collects user data has some chance to reveal that data to people who shouldn't have access to it, even if only through hacking. A security flaw is certain parts of the BESI-C system could allow quite sensitive data to be revealed. If the monitoring system were compromised, an attacker could gain access to live data about the state of patients and their homes. This data could allow someone to determine when patients are away from home or asleep. If an attacker gained complete control over the wall-mounted sensor packs, they could gain even more detailed data. They would have access to the same data exposed by the monitoring system, as well as access to microphones located on the sensor packs. It is therefore imperative for the BESI-C system to take security seriously. There is a certain level of security risk inherent to any system, especially one that leaves computers available for people to physically access, but the BESI-C project must secure the system in every way possible.

Ethical concerns over patient autonomy were also present in multiple areas of research that were considered and might occur in the BESI-C system. Each of these systems allows the user to relinquish control of some aspect of their lives to a system which can manage it on their behalf. This can be very helpful to someone if they choose to relinquish this control and it prevents them from worrying about that issue, but creates serious ethical concerns if forced on someone. This means that such systems will naturally be more useful to some than others, and that will likely also be the case with the BESI-C system. Some users will likely love not having to deal with timing their own medications, but others will be uncomfortable with having computers controlling something so essential to their comfort. Managing issues of patient autonomy is an immensely complex issue and must be considered throughout the development of the BESI-C system. It is essential that the team of engineers and nurses who work on the BESI-C project listen to users and potential users of the system. Their concerns about the system should be addressed whenever possible and designers must strive to create a system which will be useful, without burdening users.

The final major ethical concern about BESI-C I found was the threat of algorithmic bias. The machine learning algorithm that is planned to provide the BESI-C system's core functionality, so a flaw or bias in that algorithm could create serious issues. It is obvious that this algorithm needs to be very thoroughly tested and evaluated before patients use it, but more that thorough testing is required to avoid algorithmic bias. It will be essential to collect a diverse sample of training data which properly represents all groups that will use the system. Such large scale data collection is not possible with the limited number of prototypes the BESI-C project currently has access to, but will be necessary in the future. When the final training dataset is assembled it must be evaluated to ensure that marginalized groups are represented in the data. Additionally, BESI-C's algorithm must be tested and evaluated for bias proactively before it is put into general use. As prototype algorithms are developed and tested by the project, their efficacy must be evaluated for specific demographics as well as for the general population. This will ensure that the final BESI-C system is useful and effective for everyone who uses it, not just patients in a majority group.

References

Chi, N.-C., & Demiris, G. (2017). Family caregivers' pain management in end-of-life care: A systematic review. *The American Journal of Hospice & Palliative Care*, *34*(5), 470–485. https://doi.org/10.1177/1049909116637359

Hern, A. (2018). Fitness tracking app strava gives away location of secret US army bases. *The Guardian*. Retrieved October 17, 2021, from https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

LeBaron, V., Bennett, R., Alam, R., Blackhall, L., Gordon, K., Hayes, J., Homdee, N., Jones, R., Martinez, Y., Ogunjirin, E., Thomas, T., & Lach, J. (2020). Understanding the experience of cancer pain from the perspective of patients and family caregivers to inform design of an in-home smart health system: Multimethod approach [Company: JMIR Formative Research Distributor: JMIR Formative Research Institution: JMIR Formative Research Label: JMIR Formative Research Publisher: JMIR Publications Inc., Toronto, Canada]. *JMIR Formative Research*, *4*(8), e20836. https://doi.org/10.2196/20836

LeBaron, V., Hayes, J., Gordon, K., Alam, R., Homdee, N., Martinez, Y., Ogunjirin, E., Thomas, T., Jones, R., Blackhall, L., & Lach, J. (2019). Leveraging smart health technology to empower patients and family caregivers in managing cancer pain: Protocol for a feasibility study [Company: JMIR Research Protocols Distributor: JMIR Research Protocols Institution: JMIR Research Protocols Label: JMIR Research Protocols Publisher: JMIR Publications Inc., Toronto, Canada]. *JMIR Research Protocols*, *8*(12), e16178. https://doi.org/10.2196/16178

McCradden, M. D., Joshi, S., Anderson, J. A., Mazwi, M., Goldenberg, A., & Zlotnik Shaul, R. (2020). Patient safety and quality improvement: Ethical principles for a regulatory approach to bias in healthcare machine learning. *Journal of the American Medical Informatics Association*, *27*(12), 2024–2027. https://doi.org/10.1093/jamia/ocaa085

Niemeyer-Guimarães, M., & Schramm, F. R. (2017). The exercise of autonomy by older cancer patients in palliative care: The biotechnoscientific and biopolitical paradigms and the bioethics of protection [Publisher: SAGE Publications Ltd STM]. *Palliative Care: Research and Treatment*, *10*, 1178224216684831. https://doi.org/10.1177/1178224216684831

O'Neil, C. (2016, September 6). *Weapons of math destruction: How big data increases inequality and threatens democracy* [Google-Books-ID: NgEwCwAAQBAJ]. Crown.

Oudshoorn, N. (2015). Sustaining cyborgs: Sensing and tuning agencies of pacemakers and implantable cardioverter defibrillators [Publisher: SAGE Publications Ltd]. *Social Studies of Science*, *45*(1), 56–76. https://doi.org/10.1177/0306312714557377

Sharon, T. (2017). Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology*, *30*(1), 93–121. https://doi.org/10.1007/s13347-016-0215-5

Silvia Gonella, Riccardo Sperlinga, Veronica Sciannameo, Valerio Dimonte, & Sara Campagna. (2019, June 21). *Characteristics of breakthrough pain and its impact on quality of life in terminally ill cancer patients - , 2019*. Retrieved February 19, 2022, from https://journals.sagepub.com/doi/10.1177/1534735419859095

Street, A. F., & Love, A. (2005). Dimensions of privacy in palliative care: Views of health professionals. *Social Science & Medicine*, *60*(8), 1795–1804. https://doi.org/10.1016/j.socscimed.2004.08.021

Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery [Publisher: Mary Ann Liebert, Inc., publishers]. *Big Data*, *1*(2), 85–99. https://doi.org/10.1089/big.2012.0002

Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S. (2021). Mitigating bias in machine learning for medicine [Number: 1 Publisher: Nature Publishing Group]. *Communications Medicine*, *1*(1), 1–3. https://doi.org/10.1038/s43856-021-00028-w

Ward, S. E., Goldberg, N., Miller-McCauley, V., Mueller, C., Nolan, A., Pawlik-Plank, D., Robbins, A., Stormoen, D., & Weissman, D. E. (1993). Patient-related barriers to management of cancer pain. *Pain*, *52*(3), 319–324. https://doi.org/10.1016/0304-3959(93)90165-L

Yaraghi, N., & Gopal, R. D. (2018). The role of HIPAA omnibus rules in reducing the frequency of medical data breaches: Insights from an empirical study [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10 0009.12314]. *The Milbank Quarterly*, *96*(1), 144–166. https://doi.org/10.1111/1468-0009.12314

Behavioral and Environmental Sensing and Intervention for Cancer
(Technical Topic)

Effects of Predictive Algorithms in Opioid Pain Management
(STS Topic)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Penn Bauman
Spring, 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature ______________________________ Date __________
Penn Bauman

Signature ______________________________ Date __________
Rosanne Vrugtman PhD, Department of Computer Science.

Signature ______________________________ Date __________
Daniel G. Graham PhD, Department of Computer Science.

Signature ______________________________ Date __________
Joshua Earle, Department of Engineering and Society

Introduction

For years opioid addiction has ravaged many communities throughout the United States, killing hundreds of thousands of people (Jones et al., 2018, p. 1). As this epidemic has become more apparent, numerous improvements have been made to our medical system to decrease the chances of opioid painkillers leading to addiction, but there are still many for whom it is a serious risk. One group of people with such a risk are cancer patients using opioids to treat chronic pain, a relatively common situation due to no alternative high strength painkillers. In these long term pain management situations, avoiding addiction requires carefully managing the frequency and dosing of medication to ensure the patient is comfortable without unnecessary drug intake. This balance can rarely be maintained perfectly, as variation in various biological systems each data causes inevitable difficulties. This leaves doctors and patients stuck between giving patients enough medication to prevent any pain, which will certainly involve excess medication, or ensuring the patient is never over medicated, making breakthrough pain inevitable. The project this research is based on attempts to reduce both breakthrough pain and overmedication by using machine learning to dosing medication based on real-time patient data.

Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) is a joint research project of the University of Virginia School of Engineering and Applied Science and Nursing School. The project is funded by the National Institutes of Health and is currently under active research and development, with initial versions of the data collection system being tested with real patients. The system aggregates environmental, biometric, and qualitative experiential data which is then utilized by a personalized machine learning algorithm to predict when pain events will appear and when medication is required to prevent them. If successful, this predictive system could allow effective pain suppression with fewer total opioids consumed than traditional dosing methods, reducing the chances of addiction.

Technical Summary

Supporting the machine learning algorithms that will drive the decision making aspects of the BESI-C project is a set of custom data collection tools. These tools, in the form of custom hardware and software, allow the system to collect quantitative environmental and biometric data, as well as qualitative data about the pain experience of the patient. For each patient this system must be set up in their home with various elements placed at strategic locations and then smartwatches must be worn by both the cancer patient and their primary caregiver throughout the data collection period.

Each deployment of the BESI-C system utilizes two android smart watches, one worn by the cancer patient and one by their primary caregiver. Each of these watches is loaded with a custom application which allows participants to record and describe pain events. Each time a patient takes a painkiller or experiences a breakthrough pain event the watches are used to record this information. The watches also collect qualitative data from both the patient and caregiver through several surveys. There are regularly scheduled surveys which ask both patient and caregiver questions about the patient's pain level, allowing for comparison between their results. Forty minutes after a breakthrough pain event, the watches will automatically ask the patient and caregiver a set of questions about what has happened since the record pain event. Additionally, these watches collect step count, heart rate, acceleration, and relative location data and transmit it to the cloud for processing.

Environmental data for the BESI-C system is collected in each patient home through four sensor packs each placed in a different room. These packs are built using a Raspberry Pi Zero W, a tiny single board computer, wired to a custom sensor board combining various sensors. When installed each pack is stuck to a wall in one of the four most used rooms in a patient's home, positioned so that the light sensors and microphones can collect accurate readings. The relays each collected four simple data points each second: the temperature, the barometric pressure, the humidity, and the light intensity. Additionally, each relay collects half second samples of sound and processes them locally to extract 32 audio features which can be used from analysis without recording patients and violating their privacy. Once this data is collected, the relay periodically uploads the data in bulk to the cloud where it can be sorted and processed.

Finally a LTE mobile hotspot, Raspberry Pi 4B mini-computer, and an android smartphone are placed out of the way in the home. The Raspberry Pi 4B is used as a central connection point of all other devices, and is then direction connected to the mobile hotspot. The mobile hotspot provides a stable internet connection which is used by the various devices in the home to upload data to and receive directives from cloud services and databases from the project. The android smartphone simply pairs to the two smartwatches used and ensures they can connect to the wider internet properly.

STS Methodology

The quantified self refers to the phenomenon of individuals choosing to regularly monitor various aspects of their lives in order to somehow optimize them (Swan, 2013, p. 1). This self tracking can be simple logging of events, such as recording sleep time, or as complex as aggregating multiple sets of self collected personal data form sources such as smartphone applications such as the iOS Health App or home medical measurement equipment. This data can then be used in a variety of ways, such as tracking health over time, correlating particular behaviors with good or bad outcomes, or finding the most healthy sleep schedule. This category of technologies and practices is still quite new and is currently performed primarily by individuals, making it difficult to thoroughly or completely study them. Despite this, research has been done into various aspects and examples of quantified self tracking and a number of areas of common areas of concern have emerged. Among these areas are issues of user privacy and autonomy, as well as the threat of over-quantification alienating people from their own experiences (Sharon, 2017, p. 106). Privacy in particular has become an issue when digital self tracking systems handle large quantities of user data, with one run tracking app revealing the locations of secret US Army bases in Syria (Hern, 2018).

While the BESI-C system does not fit the typical pattern of a quantified self systems, because it directly involves medical professionals not just patients themselves, considering the concerns generated by quantified self research can be valuable to anticipating possible social outcomes of the system. Each of the areas of concern that have emerged in research about quantified self systems are also highly applicable to the BESI-C system, although the system's official medical status changes them in important ways. By analyzing the ways BESI-C either embodies or avoids these concerns, I will outline important ethical consideration for the future development and use of the system. In the following discussion I will analyze the BESI-C projects using the major quantified self concerns of privacy and autonomy, as well as consider its addition of a machine learning to the health data feedback loop.

Discussion

A major area of concern for the BESI-C project, quantified self systems, and any system which regularly collects user data is the privacy of collected data (Barcena et al., 2014, p. 16). In this area BESI-C is generally more responsible than quantified self systems, due to its medical nature and required HIPAA compliance. The HIPAA requirements, which ensure the anonymization of all data and consent by patient for any use of it, mean data will be much more secure than will unregulated self tracking apps, but this does not mean the BESI-C project is without privacy concerns. A privacy and security concern is created by BESI-C use of custom sensor hardware and software, which could be compromised to expose patient data. If a malicious actor gained control of relays used by the system, they would have access to multiple microphones and other sensors within the patient's home massively compromising their privacy. This danger is impossible to completely eliminate but can be significantly reduced by a focus on security in the development of the BESI-C system, including regular updates to the software the system relies on and external security audits before the system is widely deployed.

The second major privacy concern presented by the BESI-C system is the quantity of data collected. While many quantified self systems can collect data quite regularly and thoroughly a

BESI-C system produces dozens of data points every second 24 hours a day. When this quantity of data is collected it can become difficult to properly anonymize data because the data itself can reveal valuable or even identify information (Ding et al., 2010). The specific data values collected by the BESI-C system may not enough for this type of de-anonymization, but if additional data collection is added it could become possible. To ensure this never becomes a possibility, any new data sources should be carefully weighed against their risk of increasing patient exposure, while also considering the system as a whole to ensure data can not be combined with other collected data to extract information usable by a malicious actor.

Another shared area of concern with quantified self systems and the BESI-C project is the autonomy of users. In the case of standard quantified self systems this concern is created by data being used to make decisions for people, but the BESI-C project adds additional layers of autonomy loss through its involvement of medical professionals and use of machine learning. While users will retain final control over their actual consumption of medication, they are likely to respond to the BESI-C systems manipulation of this process in a variety of ways. Patients are likely to respond to these new sources of control in varying ways. Some users are likely to embrace such a system, particularly if the system works well for them. A few patients may even actively involve themselves in tuning and evaluating the system, as occurs with some pacemaker recipients (Oudshoorn, 2015, p. 69). However there will be other patients who will be uncomfortable with a computer making health choices for them, especially older patients who are less comfortable with technology. As the project continues these types of responses must be taken into account, and if possible the system should be tweaked to prevent users from experiencing the system as too overbearing.

Finally the BESI-C project differs significantly from many quantified self systems by including machine learning in its system. This can make the system much simpler to user and more efficient, but also presents unique issues. In self operated quantified self systems the tight level of control a user has over how their data is used and interpreted is key these data preventing systems become misaligned with the users experience (Ajana, 2017, p. 3). This safeguard is not present in the BESI-C system, although the regular surveys given to patients help ensure that the patient experience is considered by the system. This danger will only become more present if the BESI-C project expands in scale. As the number of patients monitored by the pain prediction algorithm increases the chances that any one patient may be failed by the algorithm increase significantly. This threat can be further compounded by external societal factors which may make particular patients' data different from others. If most users of the system are from a specific group, likely upper and middle class individuals due to healthcare costs, patients who deviate from the norm may have the system behave significantly worse for them than for more typical patients (O'Neil, 2016). This issue will be mitigated somewhat by BESI-C's use of personalized models, but creating these personalizations and ensuring their efficacy will be a constant task for the project.

In summary there are a number of ways the BESI-C system could potentially produce negative outcomes for its patients. These various dangers should be monitored and considered as the system continues to develop and grow. To prevent dangerous growths in scope, any expansion of the system should be undertaken only if it provides significant benefits to patients and does not create any new vulnerabilities for them. To ensure patients are not harmed or ignored by the algorithm that will so significantly influence their lives, the system should be assessed not only on its ability to provide appropriate care, but also on whether the patient experience is improved by such a system. If the BESI-C system is to gain widespread adoption and achieve its lofty goals of helping cancer patients, it is essential that it centers these patients and provides them real value.

References

Ajana, B. (2017, February 1). *Digital health and the biopolitics of the quantified self*. Retrieved October 3, 2021, from https://journals.sagepub.com/doi/full/10.1177/2055207616689509

Barcena, M. B., Wueest, C., & Lau, H. (2014). How safe is your quantified self?, 38.

Ding, X., Zhang, L., Wan, Z., & Gu, M. (2010). A brief survey on de-anonymization attacks in online social networks. *2010 International Conference on Computational Aspects of Social Networks*, 611–615. https://doi.org/10.1109/CASoN.2010.139

Hern, A. (2018). Fitness tracking app strava gives away location of secret US army bases. *The Guardian*. Retrieved October 17, 2021, from https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

Jones, M. R., Viswanath, O., Peck, J., Kaye, A. D., Gill, J. S., & Simopoulos, T. T. (2018). A brief history of the opioid epidemic and strategies for pain medicine. *Pain and Therapy*, *7*(1), 13–21. https://doi.org/10.1007/s40122-018-0097-6

O'Neil, C. (2016, September 6). *Weapons of math destruction: How big data increases inequality and threatens democracy* [Google-Books-ID: NgEwCwAAQBAJ]. Crown.

Oudshoorn, N. (2015). Sustaining cyborgs: Sensing and tuning agencies of pacemakers and implantable cardioverter defibrillators [Publisher: SAGE Publications Ltd]. *Social Studies of Science*, *45*(1), 56–76. https://doi.org/10.1177/0306312714557377

Sharon, T. (2017). Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology*, *30*(1), 93–121. https://doi.org/10.1007/s13347-016-0215-5

Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery [Publisher: Mary Ann Liebert, Inc., publishers]. *Big Data*, *1*(2), 85–99. https://doi.org/10.1089/big.2012.0002