Proposal Regulation for Invasive BCI Testing on the Able-Bodied

Brain-computer interfaces (BCI) are direct communication pathways between a brain and an external device. Although there are a variety of types of BCIs, the ones discussed in this paper are accomplished through neural implants, which are devices implanted in the brain that can interact with neurons through electricity by firing electric pulse patterns. They often utilize artificial intelligence algorithms – capable of recognizing very complex patterns in data – to process ("reading") complex neural data and to make a decision ("writing"). Because they are implanted in the brain, rather than other "noninvasive" BCI methods that place electrodes on the surface of the skull, neural-implant BCIs can collect much higher quality signaling data.

Companies like Neuralink and Kernel, with the support of technology giants like Facebook and Microsoft[1], are seeking to connect artificial intelligence with the brain, ultimately with the goal of enhanced human cognition. For instance, current first steps include attempting to aid the memory of humans by "reading" and "writing" them to computer chips, such that thoughts can be stored and accessed more easily[2]. As entrepreneur and venture capitalist Bryan Johnson – founder of Kernel – writes, the goal is to "program... our neural code... [to] enable us to author ourselves and our existence in ways that were previously unimaginable."[3] The Defense Advanced Research Projects Agency, a U.S. research and development agency, is also developing BCIs for able-bodied service members with the intent of improving control of unmanned aerial vehicles and defense systems, or successful multitasking in complex military missions[4]. Similarly, Neuralink – founded by Elon Musk to produce BCIs – is looking forward to clinical trials[5], eventually with the vision of enabling machine-aided human intelligence.

Part of these organizations' efforts is joining a larger existing body of research around manipulating the brain to address a variety of problems. Neural implants have been used since their first approval by the FDA in 1997 to treat many brain-related disorders, like Tourette syndrome or Parkinson’s disease. Brain-computer interfaces can allow those that are paralyzed or that have had amputations to manipulate robotic prosthetics. By manipulating nerves that link the brain stem to most key organs, researchers can also “hack” the body to treat heart failure, stroke, and other non-neural ailments[6]. What makes organizations like Kernel and Neuralink unique, though, is their push towards general human testing on able-bodied people.

In 2019, the FDA released initial draft guidance on BCI testing for the disabled[7]. It contained broad guidelines for testing, like that the risks should not outweigh the expected benefits and completion of a risk analysis. However, the FDA hasn’t come to a BCI policy on testing on able-bodied people. Given that BCI developers are moving to do so, it is imperative to develop regulation for this soon. This paper will argue for extending the current FDA draft guidance about BCI testing to include able-bodied testing by offering three important points of regulation. These points will be justified and evaluated on the FDA’s published safety principle that “the risks to the subjects … [should] not [be] outweighed by the anticipated benefits to the subjects and the importance of the knowledge to be gained.”[8]

Invasive BCIs are associated with significant risk, for several reasons. Since they need to be placed deep in the head, even if done through supposedly “minimally invasive” methods like Neuralink’s surgery robot, there is always a substantial risk with surgery that opens the skull to expose the brain. A more pressing risk, though, is the deterioration of BCI quality. With time, invasive BCIs are prone to scar-tissue build-up, which can cause weak or even completely lost signals[9]. To combat this, higher electrical current levels are required, which can cause tissue damage and overstimulate unintended regions[10]. Alternatively, even riskier invasive surgery can be performed routinely to switch electrodes[11]. These dangers are compounded by complications in the removal of invasive BCIs. Neural lace, the form of neural implant that is being used by Neuralink and many other BCI developers, is an ultra-thin mesh with interspersed electrodes[12]. Because the brain grows through this mesh by design, removing the implant safely is, in all practicality, impossible[13]. Because of all these and more dangers associated with BCIs, three researchers concluded in a study that, "implantable brain-machine interfaces … are ethically unsound in all but a handful of rare cases"[14].

Meanwhile, through continued extensive research, the quality of older EEG recording and newer magnetic methods (among others) – all noninvasive and safer – are rising to become comparable in recording quality to invasive BCIs[15] [16]. The current draft guidance regarding the sustainability and safety of the BCI requires only a verification that the BCI functions as intended over long durations at a minimum risk. Although current guidance is vague about what risk level is acceptable as "minimum", using the principle of comparing risk and benefit, evidence that the invasive BCI collects data of significantly better quality than recent non-invasive recording methods should need to be provided. It should be shown how this increased quality results in an improvement in performance for the task the data is used for, like more precise decision-making that could not be made on noninvasively collected data. This requirement eliminates the unnecessary risk posed by invasive BCIs that do not provide a meaningful increased benefit to its function when a lower-risk method could be used. This solution optimizes the information gained and benefit to BCI functionality relative to the associated risk of obtaining that information.

The security of an implanted BCI poses another significant concern. In the FDA's clinical trial considerations for brain-computer interface devices, despite software's integral role in BCIs, only one of forty-two pages is designated to discussing it[7]. Cybersecurity standards for BCIs are the same as those across all medical devices, despite the unique danger associated with invasive BCIs being embedded in the brain. A 2018 Belgian team of researchers found that wireless neural implanted devices could be remotely monitored or controlled such that a person's neural data – which includes their emotions, or even their thoughts – can be read[17]. Furthermore, undesired electrical signals could be remotely delivered to the victim's brain, which could cause paralysis or death. Another group of researchers performed software attacks on BCIs and found that thoughts containing information like PINs, addresses, and other information could be partially decrypted, posing privacy concerns[18].

The FDA's general cybersecurity standards – for which BCIs are held to – identify principles, but not security principles. For instance, while the standards outline principles that cybersecurity functions to identify and protect against threats, to limit access to trusted users only, etc., they do not consider *how* such threats should be protected from, only that their performance in testing be adequate. On the other hand, security protocols are recognized algorithmic frameworks or methods to secure the transmission of data – they provide a structure to control *how* the system is secured[19]. Given that many developers of these BCIs plan to do so commercially and on a broader scale, not controlling the processes of cybersecurity could lead to dangerous outcomes. Computer viruses "evolve" to exploit weaknesses in older software[20]; hence, despite almost all computers being built with a basic form of malware protection, 30% of U.S. computers are infected with some form of malware[21]. The lesson to learn is that scattered and unorganized malware protection across devices – even protections that "pass a test" for performance – results in exploitation. Since BCIs are computers linked to the brain, malware poses much higher a risk; a malware rate as high as 30% warrants concern for the security of

BCIs. Hence, the FDA should regulate how systems are secured by requiring BCI developers to test and update the same security protocol. To derive these protocols, the FDA and other institutions will need to invest in research to identify and continually update security protocols that protect against BCI threats.

Another software problem is the difficulty in regulating adaptive algorithms that learn over time (artificial intelligence), which are an essential important part of BCIs. There are two concerns: firstly, as models adapt to a specific patient, they become "new versions" that have not been tested. Secondly, recent research shows that the massive quantity of parameters in these intelligent adaptive algorithms leads to the problem of *underspecification,* in which an algorithm can perform well on testing data but perform poorly when it is deployed[22]. In the search space for solutions, there is usually only one true solution, but also many "artificial solutions" that can optimize the metric without "understanding" the phenomena being modeled, resulting in underspecification. The FDA released a 2019 whitepaper concerning the regulation of these adaptive algorithms, which suggests that learning updates must be confirmed by the FDA before they are implemented, addressing the first concern[23]. In January of 2021, the FDA released an action plan for adaptive algorithms[24], but it is geared more towards passive applications of AI, like software that uses AI to guide the user in manipulating ultrasound probes to produce the highest-quality images[25]. Neither the 2019 whitepaper nor the 2021 action plan addressed the problem of underspecification. However, the frequent, active, and high-impact usage of AI in invasive BCIs, which continuously listen and speak to the brain, are especially prone to underspecification. The FDA needs to work towards addressing this problem in its approval for able-bodied testing.

These concerns are less important in the context of using invasive BCIs to aid the disabled because they stand to gain more than they lose: restored functionality of a body part or treatment of a disorder often outweighs the risk of software malfunction or even invasive surgery. This paper does not argue that these changes should be implemented for BCI testing on the disabled, but instead as considerations for BCI testing policy on the able-bodied, given that this group has more to lose from the concerns discussed.

On the FDA's safety principle of risk and gain, this paper argued for addition to FDA invasive BCI policy for testing on able-bodied people in three respects. Firstly, BCI developers must not only demonstrate existing safety standards but must also demonstrate significant improvement in quality of data from recent noninvasive recording methods needed for functionality. Secondly, the FDA needs to develop specific security protocols and require BCI developers to implement and update those protocols, rather than outlining broad cybersecurity principles. Thirdly, FDA regulation for adaptive algorithms needs to update its action plan to include consideration of underspecification. As research into this recently discovered problem is limited, initial steps should be pursued in the form of drafting and discussion with the deep learning community. BCI developers are pushing boundaries at the intersection of biology and technology; regulation must catch up to ensure it is done so safe and responsibly.

Citations

[1] Matthew Sample et al. Brain-computer interfaces and personhood: interdisciplinary deliberations on neural technology. Journal of Neural Engineering. 2019 [cited 8 February 2021]. Available from: iopscience.iop.org/article/10.1088/1741-2552/ab39cd/pdf.

[2] Brain Implants that Augment the Human Brain Using AI [Internet]. Nanalyze; 2016 [cited 8 February 2021]. Available from: www.nanalyze.com/2016/11/brain-implants-ai-kernel/.

[3] Bryan Johnson. Kernel's Quest to Enhance Human Intelligence [Internet]. Medium; 2016 [cited 8 February 2021]. Available from: bryan-johnson.medium.com/kernels-quest-to-enhance-human-intelligence-7da5e16fa16c#.l97sv0qnd.

[4] Al Emondi. Next-Generation Nonsurgical Neurotechnology [Internet]. Defense Advanced Research Projects Agency; [cited 8 February 2021]. Available from: www.darpa.mil/program/next-generation-nonsurgical-neurotechnology.

[5] Rebecca Robbins, Erin Brodwin. Elon Musk's Neuralink unveils prototype of brain implants – and looks toward clinical trials [Internet]. Stat; 2020 [cited 8 February 2021]. Available from: www.statnews.com/2020/08/28/elon-musks-neuralink-unveils-prototype-of-brain-implants-and-looks-toward-clinical-trials/.

[6] Emily Waltz. How Do Neural Implants Work? [Internet]. IEEE Spectrum; 2020 [cited 8 February 2021]. Available from: spectrum.ieee.org/the-human-os/biomedical/devices/what-is-neural-implant-neuromodulation-brain-implants-electroceuticals-neuralink-definition-examples.

[7] Implanted Brain-Computer Interface (BCI) Devices for Patients with Paralysis of Amputation – Non-clinical Testing and Clinical Considerations [Internet]. U.S. Food and Drug Administration; 2019 [cited 8 February 2021]. Available from: www.fda.gov/media/120362/download

[8] Michael Hoffman. FDA Regulatory Process [Internet]. U.S. Food and Drug Administration; [cited 8 February 2021]. Available from: http://www.fda.gov/media/90419/download.

[9] Zuana Koudelkova, Martin Strmiska. Introduction to the identification of brain waves based on their frequency. MATEC Web of Conferences. 2018 [cited 8 February 2021]. Available from: www.researchgate.net/publication/328097315_Introduction_to_the_identification_of_brain_waves_based_on_their_frequency.

[10] Eran Klein, Jeffrey Ojemann. Informed consent in implantable BCI research: identification of research risks and recommendations for development of best practices. Journal of Neural Engineering. 2016 [cited 8 February 2021]. Available from: iopscience.iop.org/article/10.1088/1741-2560/13/4/043001/pdf.

[11] Steven C. Mcgie, Mary K. Nagai, Tania Artinian-Shaheen. Clinical Ethical Concerns in the Implantation of Brain-Machine Interfaces. IEEE Pulse. 2013 [cited 8 February 2021]. Available from: ieeexplore-ieee-org.offcampus.lib.washington.edu/stamp/stamp.jsp?tp=&arnumber=6493503&tag=1.

[12] What is neural lace? GMI Summit. [cited 10 February 2021]. Available form: gmisummit.com/wp-content/uploads/2019/03/What-is-neural-lace.pdf.

[13] So, you regret that brain implant: Now what? Inverse. [cited 10 February 2021]. Available from: www.inverse.com/article/31397-neural-lace-neuralink-brain-machine-interface-brain-implant-electrode.

[14] Steven C. Mcgie, Mary K. Nagai, Tania Artinian-Shaheen. Clinical Ethical Concerns in the Implantation of Brain-Machine Interfaces: Part II: Specific clinical and technical issues affecting ethical soundness. IEEE Pulse. 2013 [cited 8 February 2021]. Available from: ieeexplore.ieee.org/document/6493503/authors#authors.

[15] Joonas Iivanainen, Matti Stenroos, Lauri Parkkonen. Measuring MEG closer to the brain: Performance of on-scalp sensor arrays. NeuroImage. 2017 [cited 8 February 2021]. Available from: www.sciencedirect.com/science/article/pii/S1053811916307704.

[16] Jordy Thielen, Philip van den Broek, Jason Farquhar, Peter Desain. Broad-Band Visually Evoked Potentials: Re(con)volution in Brain-Computer Interfacing. Plos One. 2015 [cited 8 February 2021]. Available from: doi.org/10.1371/journal.pone.0133797.

[17] Eduard Marin et al. Securing Wireless Neurostimulators. Association for Computing Machinery. 2018 [cited 10 February 2021]. Available from: www.esat.kuleuven.be/cosic/publications/article-2803.pdf.

[18] Ivan Martinovic et al. On the Feasibility of Side-Channel Attacks with Brain-Computer Interfaces. Usenix conference. 2018 [cited 10 February 2021]. Available from: www.usenix.org/system/files/conference/usenixsecurity12/sec12-final56.pdf.

[19] Definition of security protocol. PCmag. 2021 [cited 10 February 2021]. Available from: www.pcmag.com/encyclopedia/term/security-protocol.

[20] Carey Nachenberg. The Evolving Virus Threat. Symantec. 2015 [cited 10 February 2021]. Available from: wenke.gtisc.gatech.edu/worms-readings/Nachenberg_EvolvingVirusThreat.pdf.

[21] Bojan Jovanovic. Malware statistics. Dataprot. 2019 [cited 10 February 2021]. Available from: dataprot.net/statistics/malware-statistics/.

[22] D'Amour et al. Underspecification Presents Challenges for Credibility in Modern Machine Learning. Arxiv. 2020 [cited 8 February 2021]. Available from: arxiv.org/pdf/2011.03395.pdf.

[23] Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD). U.S. Food and Drug Administration Center for Devices and Radiological Health. 2019 [cited 9 February 2021]. Available from: www.fda.gov/media/122535/download.

[24] Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. U.S. Food and Drug Administration Center for Devices and Radiological Health. 2021 [cited 9 February 2021]. Available from: www.fda.gov/media/145022/download.

[25] FDA Authorizes Marketing of First Cardiac Ultrasound Software That Uses Artificial Intelligence to Guide User. U.S. Food and Drug Administration News Release. 2020 [cited 9 February 2021]. Available from: www.fda.gov/news-events/press-announcements/fda-authorizes-marketing-first-cardiac-ultrasound-software-uses-artificial-intelligence-guide-user.