Sl. No.	Device Name [1]	Physiological Panel	Device Functionality	User	Type of ML algo used	Type of data processed	Known attacks (Attack type)	Potential impact of misprediction
1	CardioLogs ECG Analysis Platform†	Cardiovascular	Cardiac arrhythmia detector	Medical practitioners	Deep Neural Network (DNN)	Image	Chen et al. [2] I	Wrong treatment (Fatal)
2	Oxehealth Vital Signs†	Cardiovascular	Camera-based monitor for heart, pulse, and respiratory rate	Medical practitioners	Hybrid convolutional Long short term memory networks (LSTM)	Video	Albattah et al. [3] T,I	Wrong treatment
3	GI Genius‡	Gastroenterology/ Urology	Gastro-intestinal lesion detection	Medical practitioners	Convolutional neural networks (CNN) *	Video	Amin et al. [28] T	Wrong diagnosis
4	SOZO‡	Gastroenterology/ Urology	Body fluid analyzer for assessing protein-calorie malnutrition	Medical practitioners	CNN *	Numeric	Byra et al. [29] I	Wrong diagnosis
5	WellDoc BlueStar†	General hospital	Diabetes management	Medical practitioners, patients	Darknet-53 CNN	Numeric	Lal et al. [4] I	Wrong diagnosis
6	d-Nav System†	General hospital	Insulin dose predictor	Medical practitioners, patients	Multi-layer perception (MLP) and LSTM	Numeric	Zhou et al. [30] I	Wrong treatment (Fatal)
7	MBT-CA System‡	Microbiology	Spectometry	Medical practitioners	DNN *	Numeric	Meiseles et al. [5] I	Wrong diagnosis (Fatal)
8	KIDScore D3†	Obstetrics & Gynaecology	Embryo image assessment	Medical practitioners	Decentralized federated learning	Image	Nguyen et al. [31] P	Wrong diagnosis
9	NuVasive Pulse System‡	Orthopedic	Neurological monitoring	Medical practitioners	CNN *	Image	Joel et al. [6] I	Mistake in surgery (Fatal)
10	ABMD Software†	Radiology	Bone densitometer	Medical practitioners	Inception-v3 and Densenet-121 *	Image	Bortsova et al. [7] I	Wrong diagnosis
11	Deep Learning Image Reconstruction†	Radiology	X-ray reconstruction	Medical practitioners	ResNet-18	Image	Menon et al. [8] T Paul et al. [32] I	Wrong diagnosis
12	Air Next‡	Anesthesiology	Spirometer	Medical practitioners	CatBoost ResNet-50 *	Image	Vargas et al. [9] I	Wrong diagnosis
13	One Drop Blood Glucose Monitoring System‡	Clinical Chemistry	Diabetes management	Patients	MLP	Numeric	Levy-Loboda et al. [10] I	Wrong treatment (Fatal)
14	OTIS 2.1 and THiA Optical Coherence Tomography System‡	General and Plastic Surgery	Human tissue imaging	Medical practitioners	Support Vector Machines (SVM)	Image	Ma et al. [16] I	Wrong diagnosis
15	EarliPoint System‡	Neurology	Diagnosis of Pediatric Autism Spectrum Disorder	Medical practitioners	Graph Neural Network (GNN)	Image	Chen et al. [11] T	Wrong diagnosis
16	BrainScope TBI‡	Neurology	Brain injury assessment	Medical practitioners	CNN + Recurrent neural networks (RNN)	Numeric	Yu et al. [12] I	Wrong treatment (Fatal)
17	IDx-DR v2.3†	Ophthalmic	Diabetic Retinopathy Detection	Medical practitioners	Federated learning	Image	Nielsen et al. [13] I	Wrong diagnosis (loss of vision)
18	Iris Intelligent Retinal Imaging System†	Ophthalmic	Storage, management and display of retinal images	Medical practitioners	DNN	Image	Mangaokar et al. [14] I	Wrong diagnosis (loss of vision)
19	Paige Prostate†	Pathology	Cancer diagnosis	Medical practitioners	CNN	Numeric	Hu et al. [15] T	Wrong treatment (Fatal)
20	Tissue of Origin Test Kit‡	Pathology	Malignant Tumor diagnosis	Medical practitioners	SVM	Image	Ma et al. [16] I	Wrong treatment (Fatal)

TABLE I: A study of different FDA-Approved ML-enabled medical devices and their security vulnerabilities
†: Software as medical device, ‡: Software in medical device, *: Best-guessed ML algorithm,
T: Training-time attack, I: Inference-time attack, P: Privacy attack

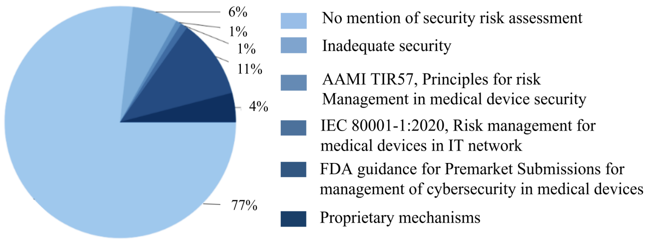

Fig. 5: Security risk assessment techniques by manufacturers of FDA-approved ML-enabled medical systems based on [39]

existing risk assessment techniques, which as we have seen, are not sufficient for assessing the severity of the security risks in ML-enabled connected medical systems. Therefore, risk assessment of ML-enabled medical devices remains an open challenge. Developing such a risk assessment technique would require bridging the domains of cybersecurity and medicine.

VI. Conclusion and Future Research Directions

We presented a detailed study of security risks associated with modern AI/ML-enabled medical devices, mainly due to vulnerabilities in connected peripheral devices. We conducted a systematic security analysis of commercial AI/ML-enabled devices that were approved by the U.S. Food and Drug Administration (FDA). Our analysis shows that many of these devices are vulnerable to existing adversarial attacks, which raises concerns about the appropriateness of using such safety-critical devices on patients in the real world. We validate our analysis by performing a case study, where we execute a realistic adversarial attack on an ML-enabled blood glucose monitoring system. Through this case study, we identify security risks in the glucose monitoring system. Further, we studied state-of-the-art risk assessment frameworks to highlight their drawbacks in identifying security risks in connected ML-enabled medical systems, and the need for a new framework.

Our work opens up three interesting future work directions – **(1)** Automated risk identification: Automating the risk identification process at scale would benefit device manufacturers as well as the security research community. This would require identifying relevant documents on the web and parsing a huge volume of unstructured documents, while at the same time being able to relate various ML concepts; **(2)** Building personalized spatial and temporal risk profiles per patient: Our case study shows that attacks on ML-enabled medical systems cause more damage to certain patients than others. Moreover, a patient is not equally vulnerable at all points of time. An interesting research problem is to study patients' data in more detail to develop customized spatial and temporal risk profiles for every patient; and, **(3)** Efficient risk mitigation techniques: This involves designing attack-resilient ML models, determining accountable entity and enforcing accountability in risk mitigation, accounting for the costs and deployment scenario.

NOTE: We will make our code and datasets publicly available if accepted

References

[1] U.S. Food & Drug Administration. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. URL: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices, Last accessed: Nov 30, 2023.

[2] Huangxun Chen, Chenyu Huang, Qianyi Huang, Qian Zhang, and Wei Wang. Ecgadv: Generating adversarial electrocardiogram to misguide arrhythmia classification system. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3446–3453, 2020.

[3] Albatul Albattah and Murad A Rassam. Detection of adversarial attacks against the hybrid convolutional long short-term memory deep learning technique for healthcare monitoring applications. *Applied Sciences*, 13(11):6807, 2023.

[4] Sheeba Lal, Saeed Ur Rehman, Jamal Hussain Shah, Talha Meraj, Hafiz Tayyab Rauf, Robertas Damaševičius, Mazin Abed Mohammed, and Karrar Hameed Abdulkareem. Adversarial attack and defence through adversarial training and feature fusion for diabetic retinopathy recognition. *Sensors*, 21(11):3922, 2021.

[5] Amiel Meiseles, Ishai Rosenberg, Yair Motro, Lior Rokach, and Jacob Moran-Gilad. Adversarial vulnerability of deep learning models in analyzing next generation sequencing data. In *2020 IEEE International Conference on BIBM*, pages 464–468, 2020.

[6] Marina Z Joel, Sachin Umrao, Enoch Chang, Rachel Choi, Daniel Yang, James Duncan, Antonio Omuro, Roy Herbst, Harlan M Krumholz, Sanjay Aneja, et al. Adversarial attack vulnerability of deep learning models for oncologic images. *MedRxiv*, 2021.

[7] Gerda Bortsova, Cristina González-Gonzalo, Suzanne C Wetstein, Florian Dubost, Ioannis Katramados, Laurens Hogeweg, Bart Liefers, Bram van Ginneken, Josien PW Pluim, Mitko Veta, et al. Adversarial attack vulnerability of medical image analysis systems: Unexplored factors. *Medical Image Analysis*, 73:102141, 2021.

[8] Karthika Menon, V Khushi Bohra, Lakshana Murugan, Kavya Jaganathan, and Chamundeswari Arumugam. Covid-19 diagnosis from chest x-ray images using convolutional neural networks and effects of data poisoning. In *Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part IX 21*, pages 508–521. Springer, 2021.

[9] Danilo Vasconcellos Vargas and Jiawei Su. Understanding the one-pixel attack: Propagation maps and locality analysis. In *CEUR Workshop Proceedings*, volume 2640. CEUR-WS, 2020.

[10] Tamar Levy-Loboda, Eitam Sheetrit, Idit F Liberty, Alon Haim, and Nir Nissim. Personalized insulin dose manipulation attack and its detection using interval-based temporal patterns and machine learning algorithms. *Journal of Biomedical Informatics*, 132:104129, 2022.

[11] Yuzhong Chen, Jiadong Yan, Mingxin Jiang, Tuo Zhang, Zhongbo Zhao, Weihua Zhao, Jian Zheng, Dezhong Yao, Rong Zhang, Keith M Kendrick, et al. Adversarial learning based node-edge graph attention networks for autism spectrum disorder identification. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[12] Jianfeng Yu, Kai Qiu, Pengju Wang, Caixia Su, Yufeng Fan, and Yongfeng Cao. Perturbing BEAMs: EEG adversarial attack to deep learning models for epilepsy diagnosing. *BMC Medical Informatics and Decision Making*, 23(1):115, 2023.

[13] Christopher Nielsen, Anup Tuladhar, and Nils D Forkert. Investigating the vulnerability of federated learning-based diabetic retinopathy grade classification to gradient inversion attacks. In *International Workshop on Ophthalmic Medical Image Analysis*, pages 183–192. Springer, 2022.

[14] Neal Mangaokar, Jiameng Pu, Parantapa Bhattacharya, Chandan K Reddy, and Bimal Viswanath. Jekyll: Attacking medical image diagnostics using deep generative models. In *2020 IEEE EuroS&P*, pages 139–157. IEEE, 2020.

[15] Lei Hu, Da-Wei Zhou, Xiang-Yu Guo, Wen-Hao Xu, Li-Ming Wei, and Jun-Gong Zhao. Adversarial training for prostate cancer classification using magnetic resonance imaging. *Quantitative Imaging in Medicine and Surgery*, 12(6):3276, 2022.

[16] Xingjun Ma, Yuhao Niu, Lin Gu, Yisen Wang, Yitian Zhao, James Bailey, and Feng Lu. Understanding adversarial attacks on deep learning based medical image analysis systems. *Pattern Recognition*, 110:107332, 2021.

[17] U.S. Food & Drug Administration. DreaMed Advisor Pro. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/denovo.cfm?id=DEN170043, Last accessed: Nov 30, 2023.

[18] Michael Kahn. Diabetes. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5T59G.

[19] MIchelle Mouri and Madhu Badireddy. Hyperglycemia. URL: https://www.ncbi.nlm.nih.gov/books/NBK430900/, Last accessed: Nov 30, 2023.

[20] Harry Rubin-Falcone, Ian Fox, and Jenna Wiens. Deep Residual Time-Series Forecasting: Application to Blood Glucose Prediction. In *KDH@ ECAI*, pages 105–109, 2020.

[21] Kasper Rasmussen. BLURtooth: Exploiting Cross- Transport Key Derivation in Bluetooth Classic and Bluetooth Low Energy. In *AsiaCCS*. ASIA Conference on Computer and Communications Security, 2022.

[22] Dreamed Diabetes Ltd. DreaMed Advisor Pro: Manual For Personal Use iOS. URL: https://dreamed-diabetes.com/wp-content/uploads/2019/06/Dreamed-Advisor-Pro-iOS-Patient-IFU.pdf, Last accessed: Nov 30, 2023.

[23] Guillaume Dupont, Daniel Ricardo dos Santos, Elisa Costante, Jerry Den Hartog, and Sandro Etalle. A matter of life and death: analyzing the security of healthcare networks. In *SEC 2020: ICT Systems Security and Privacy Protection*, pages 355–369. Springer, 2020.

[24] SMJ Mortazavi, M Gholampour, M Haghani, G Mortazavi, and AR Mortazavi. Electromagnetic radiofrequency radiation emitted from GSM mobile phones decreases the accuracy of home blood glucose monitors. *Journal of Biomedical Physics & Engineering*, 4(3):111, 2014.

[25] From big data to precision medicine, author=Hulsen, Tim and Jamuar, Saumya S and Moody, Alan R and Karnes, Jason H and Varga, Orsolya and Hedensted, Stine and Spreafico, Roberto and Hafler, David A and McKinney, Eoin F. *Frontiers in medicine*, 6:34, 2019.

[26] Spiros V Georgakopoulos, Dimitris K Iakovidis, Michael Vasilakakis, Vassilis P Plagianakos, and Anastasios Koulaouzidis. Weakly-supervised convolutional learning for detection of inflammatory gastrointestinal lesions. In *2016 IEEE International conference on Imaging Systems and Techniques (IST)*, pages 510–514. IEEE, 2016.

[27] Amalia M Issa. Personalized medicine and the practice of medicine in the 21st century. *McGill Journal of Medicine: MJM*, 10(1):53, 2007.

[28] Muhammad Shahid Amin, Jamal Hussain Shah, Mussarat Yasmin, Ghulam Jillani Ansari, Muhamamd Attique Khan, Usman Tariq, Ye Jin Kim, and Byoungchol Chang. A two stream fusion assisted deep learning framework for stomach diseases classification. *CMC-Comput. Mater. Contin*, 73:4423–4439, 2022.

[29] Michal Byra, Grzegorz Styczynski, Cezary Szmigielski, Piotr Kalinowski, Lukasz Michalowski, Rafal Paluszkiewicz, Bogna Ziarkiewicz-Wroblewska, Krzysztof Zieniewicz, and Andrzej Nowicki. Adversarial attacks on deep learning models for fatty liver disease classification by modification of ultrasound image reconstruction method. In *2020 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4, 2020.

[30] Xugui Zhou, Maxfield Kouzel, and Homa Alemzadeh. Robustness testing of data and knowledge driven anomaly detection in cyber-physical systems. In *2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, pages 44–51. IEEE, 2022.

[31] TV Nguyen, MA Dakka, SM Diakiw, MD VerMilyea, M Perugini, JMM Hall, and D Perugini. A novel decentralized federated learning approach to train on globally distributed, poor quality, and protected private medical data. *Scientific Reports*, 12(1):8888, 2022.

[32] Rahul Paul, Matthew Schabath, Robert Gillies, Lawrence Hall, and Dmitry Goldgof. Mitigating adversarial attacks on medical image understanding systems. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 1517–1521. IEEE, 2020.

[33] Mitre. Common Vulnerabilities and Exposures. URL: https://cve.mitre.org/, Last accessed: Nov 30, 2023.

[34] Information Technology Laboratory, USA. National Vulnerability Database. URL: https://nvd.nist.gov/vuln/search, Last accessed: Nov 30, 2023.

[35] Mitre. Conexus Telemetry Protocol vulnerability. URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-6538, Last accessed: Nov 30, 2023.

[36] Wei Wang, Yao Yao, Xin Liu, Xiang Li, Pei Hao, and Ting Zhu. I can see the light: Attacks on autonomous vehicles using invisible lights. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 1930–1944, 2021.

[37] Mitre. Sony IPELA E Series Camera vulnerability (1). URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-3938, Last accessed: Nov 30, 2023.

[38] Mitre. Windows 7 vulnerability (2). URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-5921, Last accessed: Nov 30, 2023.

[39] U.S. Food & Drug Administration. Content of Premarket Submissions for Management of Cybersecurity in Medical Devices. URL: https://www.fda.gov/media/86174/download, Last accessed: Nov 30, 2023.

[40] Mitre. Shekar Endoscope vulnerability (1). URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2017-10722, Last accessed: Nov 30, 2023.

[41] Mitre. GE Healthcare Discovery vulnerability. URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2014-7232, Last accessed: Nov 30, 2023.

[42] AAMI. ANSI/AAMI/ISO 14971: 2007/(R) 2010, Medical devices—Application of risk management to medical devices.

[43] International Electrotechnical Commission et al. Iec 62304: 2006/a1: 2015. *Medical device software-Software life-cycle processes*, 2015.

[44] Mitre. H264WebCam vulnerability. URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2010-2349, Last accessed: Nov 30, 2023.

[45] U.S. Food & Drug Administration. IDx-DR v2.3. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K213037, Last accessed: Nov 30, 2023.

[46] George J Annas. Hipaa regulations: a new era of medical-record privacy? *New England Journal of Medicine*, 348:1486, 2003.

[47] Mitre. Philips MRI 1.5T and MRI 3T vulnerability (1). URL: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-26262, Last accessed: Nov 30, 2023.

[48] Kevin Eykholt, Taesung Lee, Douglas Schales, Jiyong Jang, and Ian Molloy. URET: Universal Robustness Evaluation Toolkit (for Evasion). In *USENIX Security 23*, pages 3817–3833, Anaheim, CA, August 2023. USENIX Association.

[49] U.S. Food & Drug Administration. d-Nav System. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K181916, Last accessed: Nov 30, 2023.

[50] Cindy Marling and Razvan Bunescu. The OhioT1DM dataset for blood glucose level prediction: Update 2020. In *CEUR workshop proceedings*, volume 2675, page 71. NIH Public Access, 2020.

[51] Battista Biggio, Igino Corona, Davide Maiorca, Blaine Nelson, Nedim Šrndić, Pavel Laskov, Giorgio Giacinto, and Fabio Roli. Evasion attacks against machine learning at test time. In *ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III 13*, pages 387–402. Springer, 2013.

[52] Samuel G Finlayson, John D Bowers, Joichi Ito, Jonathan L Zittrain, Andrew L Beam, and Isaac S Kohane. Adversarial attacks on medical machine learning. *Science*, 363(6433):1287–1289, 2019.

[53] Samuel G Finlayson, Hyung Won Chung, Isaac S Kohane, and Andrew L Beam. Adversarial attacks against medical deep learning systems. *arXiv preprint arXiv:1804.05296*, 2018.

[54] Nicholas Carlini and David Wagner. Towards evaluating the robustness of neural networks. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 39–57. IEEE, 2017.

[55] J.D. Meier and Microsoft Corporation. *Improving Web Application Security: Threats and Countermeasures*. Patterns & practices. Microsoft, 2003.

[56] Hu-Chen Liu, Li-Jun Zhang, Ye-Jia Ping, and Liang Wang. Failure mode and effects analysis for proactive healthcare risk evaluation: a systematic literature review. *Journal of evaluation in clinical practice*, 26(4):1320–1337, 2020.

[57] Tom Mahler, Yuval Elovici, and Yuval Shahar. A new methodology for information security risk assessment for medical devices and its evaluation. *arXiv preprint arXiv:2002.06938*, 2020.

[58] Tahreem Yaqoob, Haider Abbas, and Narmeen Shafqat. Integrated security, safety, and privacy risk assessment framework for medical devices. *IEEE journal of biomedical and health informatics*, 24(6):1752–1761, 2019.

[59] R Abraham, D Arora, M Coles, M Eckert, M Heitman, A Manion, S Moore, S Romanowsky, K Scarfone, J Stuppi, et al. Common vulnerability scoring system v3. 0: Specification document. *First*, 2015.

[60] Jonathan Spring, Eric Hatleback, Allen Householder, Art Manion, and Deana Shick. Time to Change the CVSS? *IEEE Security & Privacy*, 19(2):74–78, 2021.

[61] Eric Wu, Kevin Wu, Roxana Daneshjou, David Ouyang, Daniel E Ho, and James Zou. How medical ai devices are evaluated: limitations and recommendations from an analysis of fda approvals. *Nature Medicine*, 27(4):582–584, 2021.