

Title: AI/ML Models to Aid in the Diagnosis of COVID-19 Illness from Forced Cough Vocalizations: Good Machine Learning Practice and Good Clinical Practices from Concept to Consumer for AI/ML Software Devices

Authors: K. Kelley, M. Kelley, S. C. Kelley, A.A. Sakara, M.A. Ramirez

Affiliations:

K. Kelley, MD, FAADM – Chief Regulatory Medical Director for RAisonance, Inc.,

M. Kelley, MSN, RN, LNC – Co-Chief for Clinical and Regulatory Affairs for RAisonance, Inc.,

S. C. Kelley, DNP, AGACNP-BC, FNP-BC – SVP Clinical Trials and Quality Improvement for RAisonance, Inc.

A. A. Sakara, NP, MSN, RN, PHRN - Co-Chief for Clinical and Regulatory Affairs for RAisonance, Inc.,

M. A. Ramirez, DO, PhD - Chief Medical Officer, Chief Scientist, Chief Strategy and Planning Officer for RAisonance, Inc

Corresponding Author: Allison A. Sakara
4800 Ben Hill Trail
Lake Wales, FL 33898
hire@high-alert.com

Dedicated to Our Respected Fellow AI/ML COVID-19 Signal Data Signature Researchers and Healthcare Decision Makers who have died of COVID-19: M. Maleki and M.E. Thornton.

Abstract:

From a comprehensive and systematic search of the relevant literature on signal data signature (SDS)-based artificial intelligence/machine learning (AI/ML) systems designed to aid in the diagnosis of COVID-19 illness, we identified the highest quality articles with statistically significant data sets for a head-to-head comparison to our own model in development. Further comparisons were made to the recently released “Good Machine Learning Practice (GMLP) for Medical Device Development: Guiding Principles” and, in conclusions, we proposed supplemental principles aimed at bringing AI/ML technologies in closer alignment GMLP and Good Clinical Practices (GCP).

1. Introduction

Artificial intelligence and machine learning (AI/ML) technologies have the potential to augment and democratize health care by providing analysis and insights into the vast amount of health-related data collected pertaining to individuals, as well as their communities. The promise of AI/ML solutions to realize these goals, however, has been hampered by the failure of both Software as Medical Device (SaMD) and Software in Medical Device (SiMD) manufacturers to employ medical device industry standards for design, documentation, testing, manufacturing and deployment across the Total Product Life Cycle (TPLC).

We prospectively planned a PRISMA 2020 systematic review of the relevant literature to demonstrate reproducibility of the published models by building each COVID-19 diagnostic software system for which sufficient details were reported and to conduct head-to-head evaluations both across the completed models and in comparison, to our own AI/ML system. As a secondary objective, our intent was to determine a literature-derived performance goal (PG) for the completed models for comparison to our device with the same intended use.

2. Background

Poor data quality leading to poorly trained AI/ML models that cannot perform in a clinical setting has been described extensively in the literature [1-4]. Inadequate sample sizes for training, validation, and testing are often secondary to the lack of availability of datasets that results in a lack of statistical significance and subsequent rejection by informed clinicians. Furthermore, authors within the current literature have not utilized a prospective plan for the total product life cycle (TPLC) of their device - a necessity for regulatory approval and deployment in the clinical world - promoting a lack of prospective objectives for intended use. The literature also demonstrated a lack of ability of the current models to meet minimum standards set forth by the U.S. Food and Drug Administration (FDA) and the World Health Organization (WHO) for safety and performance. Along with the risks of bias stemming from socioeconomic disparities in data collections, each of these gaps echo statements made in current editorials presenting the challenges of advancing AI/ML technology to aid in the diagnosis of COVID-19 [1] [4].

On 29 October 2021, FDA, CanadaHealth and the Medicines and Healthcare products Regulatory Agency (MHRA) jointly developed and published 10 guiding principles for “Good Machine Learning Practice (GMLP) for Medical Device Development: Guiding Principles,” for the TPLC of SaMDs and SiMDs [5]. The GMLP is a companion to the 13 guiding principles for Good Clinical Practices (GCP) for medical device clinical trials published by the International Conference on Harmonization (ICH) and endorsed by FDA in 1997 [6-7]. With the GCP, standards were set for the design, conduct, and reporting of clinical trials and is key to bringing new medical devices to market. This article will discuss how these principles can be applied to the TPLC to align AI/ML development with clinical trial design and endpoints.

3. Materials and Methods:

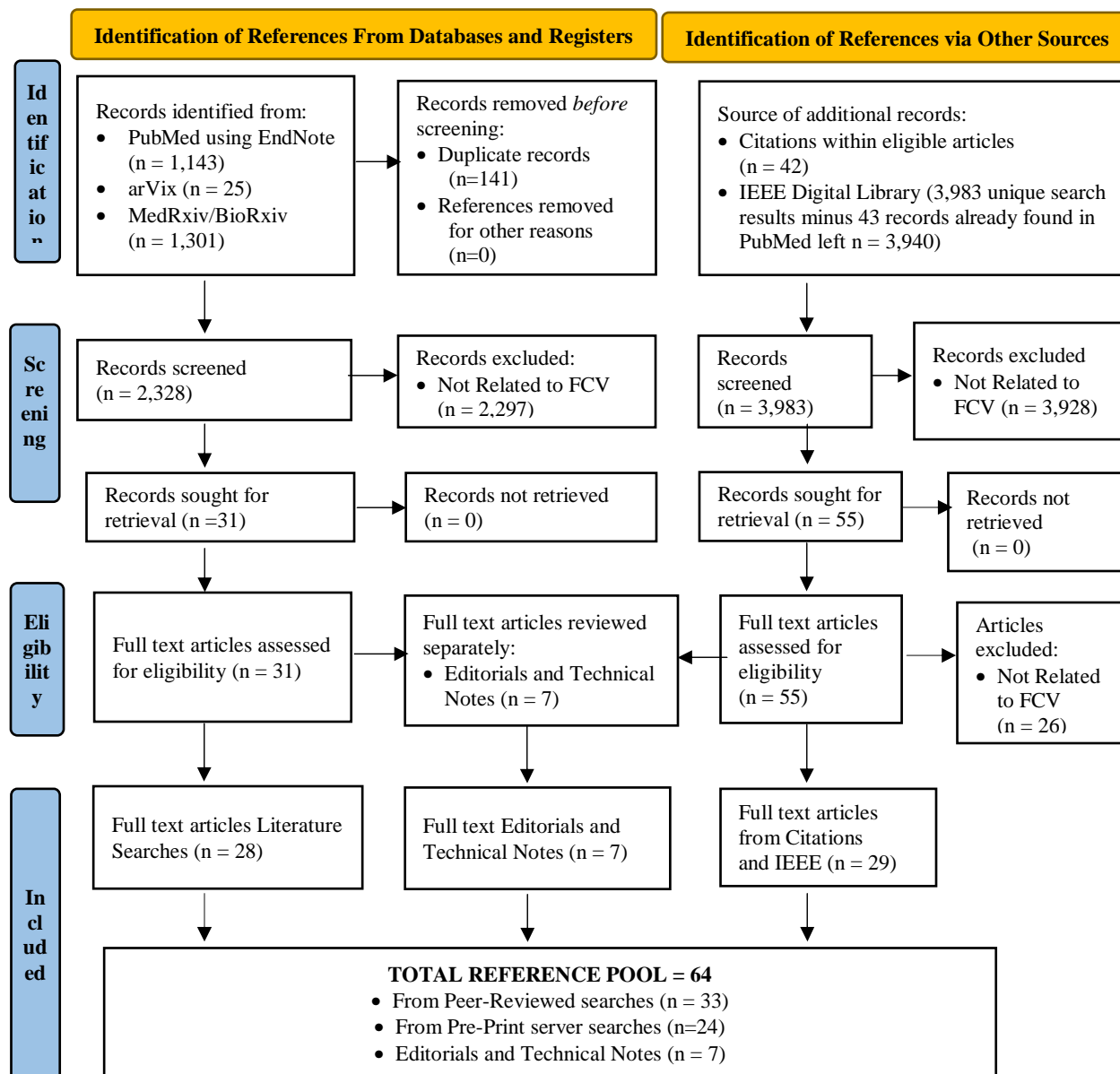
Prior reporting of our methodology and results were detailed in K. Kelley et al. [K. Kelley et. al. 2021 – Manuscript Submitted for Publication] and have been reproduced below in part for ease of reference.

As previously described by our team, we conducted systematic searches of the relevant literature on 12 October 2021 and updated on 7 November 2021, for the purpose of presenting a comparative evaluation of AI/ML systems designed to aid in the diagnosis of COVID-19 from FCV. Searches of the peer-reviewed literature were prioritized but, given the collaborative “shareware” culture of the AI/ML and Data Science communities, pre-print servers were searched for possible contributions. EndNote 2020 was the designated reference manager and PubMed was searched via this software. Serial searches of “Any Field” in PubMed, “Full Text and Metadata” in the IEEE Xplore digital library of the Institute of Electrical and Electronics Engineers, (ieeexplore.ieee.org), “All Fields” in the arXiv open-access archive (arxiv.org), and “Full Text or Abstract or Title” in bioRxiv and medRxiv (medrxiv.org) were performed using the identical search terms as listed below:

- covid and classifier
- covid and neural network
- covid and cough and artificial intelligence
- covid and cough and AI
- covid and cough and machine learning
- covid and cough and ML
- covid and cough and classifier
- cough and neural network
- forced cough vocalization

The results from these serial searches were combined and systematically filtered to achieve a final article pool from which all references would be evaluated for contribution to the stated objectives. Following a basic PRISMA 2020 workflow, our search methodology is illustrated below in Figure 1:

Figure 1 - PRISMA 2020 Flow Diagram for New Systematic Reviews (inclusive of Databases, SDS Libraries, and Other Sources)



4. Results:

From the final pool of 64 references, the remaining 33 peer-reviewed articles and 24 pre-print server records were read and evaluated in full across a multidisciplinary team of Data Scientists, Biotechnical Engineers, Healthcare Clinicians, Product Development and Clinical/Regulatory Affairs professionals. Although each of 57 references purported to include a fully-detailed AI/ML solution, only 14 contained enough information for us to attempt building the stated solution. Of those 14 references, only two peer-reviewed articles demonstrated statistical significance consistent with the requirements for a regulatory pathway.

4.1. Primary and Secondary Objectives

No single reference included sufficient details to build a complete model or system, rendering comparisons across models (including our own) unattainable. The primary objectives of the review, therefore, could not be met.

Determination of a literature-derived Performance Goal (PG) was the secondary objective of the original review and is a familiar endpoint in regulatory pathways. None of the references met all inclusion criteria, however, rendering the secondary objective unattainable.

Since the intended objectives of our original systematic review could not be met, we undertook to compare the two peer-reviewed articles [8-9] that demonstrated statistically significance data sets consistent with the requirements for a regulatory pathway and to our model in development. These comparisons will be presented in relationship to two frameworks: the principles of the recently released GMLP and the predicate principles of the GCP for Medical Devices.

4.2. GMLP Analysis

The new proposed primary objective was to compare the two articles previously identified as having Level 2a design, Moderate Grade data, and a statistically significant test set size for head-to-head comparison to each other and also to our model in development.

Table 1: Comparisons to GMLP

	Good Machine Learning Practice	Andreu-Perez et al. [8]	Verde et al. [9]	Our Model
1	<i>Multi-Disciplinary Expertise is Leveraged throughout the Total Product Life Cycle</i>	NR Author list includes 1 RN.	NR All authors are biomedical engineers or software engineers	1) Involvement of a Clinical and Regulatory Team early in planning phase of SaMD 2) Medical Advisory board to meet needs of front-line clinicians 3) Employing and consulting subject matter experts for AI/ML systems, quality management, and statistic
2	<i>Good Software Engineering and Security Practices Are Implemented</i>	1) Utilized an in-house registry that was collected in accordance with Level 2a - Moderate Grade Data 2) Bench tested several neural network architectures ultimately yielding a single, most accurate model for the final software system.	1) Utilized Coswara library that was collected in accordance with Level 2a design and Moderate Grade Data 2) Bench tested several neural network architectures ultimately yielding a single, most	1) Merged best practices for software, audio, and biotechnical engineering. 2) Clear, comprehensive documentation of software engineering practices used is key. 3) All data scientists trained in-house for biotechnical engineering certification. 4) Use of Agile software

		3) Engineered app and software to be GDPR compliant for privacy and cybersecurity	accurate model for the final software system.	development practices to support iterative nature of SaMD development. 5) Use of device and user FMEAs to promote quality 6) Establishment of cybersecurity policy and procedure.
3	<i>Clinical Study Participants and Data Sets are Representative of the Intended Patient Population</i>	Participants from both Spain and Mexico to reduce bias.	NR	Benchmark testing data sets matched to intended clinical trial site populations by: 1) Obtaining SDS registries from multiple nations. 2) Prospectively planned descriptive statistical evaluation of outcome data based upon demographic subgroups. 3) Training sets at least 2-3 times the size of the testing set. 4) Use of Exact Binomial Test to calculate training set size based on the N for the prospective testing set
4	<i>Training Data Sets Are Independent of Test Sets</i>	Data sets remained separate.	Data sets remained separate.	Data sets remained separate due to: 1) Use of SDS adjudication to evaluate all registries for quality. 2) Used of best quality data sets for training. 3) Removal of duplicate recordings within datasets.
5	<i>Selected Reference Datasets Are Based Upon Best Available Methods</i>	Prospective collection of cough sounds verified against RT-PCR.	Prospective collection of cough sounds verified against RT-PCR.	1) Find highest level of data available in existing registries 2) Use of sound adjudication to improve the quality of data 3) Prospective collection of cough sounds verified against RT-PCR and Sanger sequencing.
6	<i>Model Design is Tailored to the Available Data and Reflects the Intended Use of the</i>	Model design consistent with intended use of screening for COVID-	Model design consistent with intended use of diagnosing for	Model design consistent with intended use of diagnosing for COVID-19 illness via mobile device and has the

	<i>Device</i>	19.	COVID-19 via mobile device.	potential to diagnose other acute and chronic illnesses in the future.
7	<i>Focus Is Placed on the Performance of the Human-AI Team</i>	NR	NR	Patented HybridOps system allows the human to observe all drift and to be the gatekeeper for all development and deployment.
8	<i>Testing Demonstrates Device Performance During Clinically Relevant Conditions</i>	The machine was tested against RT-PCR and information was obtained from subjects on symptomology.	Model was tested against RT-PCR and information obtained from physicians on symptomatology.	1) Benchtop tested against RT-PCR and information obtained from physicians on symptomatology. 2) RT-PCR and Sanger sequencing for clinical trial testing to ensure true positive and negative.
9	<i>Users are Provided Clear, Essential Information</i>	1) No usability data reported. 2) Results are delivered as a probability of the FCV indicating COVID-19 illness as a screening test. The system does not give a definitive diagnosis of COVID-19.	1) No usability data reported. 2) Results are delivered as diagnostic for COVID-19. No clear description for the final app or reporting system.	1) IFU created per regulatory standards 2) Clear, concise, step-by-step instruction sheets for quick guide 3) Help function on every screen of the app 4) Instructional video available within the app 5) Prospective usability data collection during clinical trial 6) Prospective plan for reporting of results through LIMS system
10	<i>Deployed Models are Monitored for Performance and Re-training Risks are Managed</i>	NR	NR	Hybrid OPS provides a three-channel deployment and development environment that monitors for context drift, model drift and model performance compared to the originally approved system, reducing bias.

NR = Not Reported

4.3. GCP Analysis

Once an AI/ML medical device is developed and ready for clinical testing, there are 13 Good Clinical Practice (GCP) principles that must be observed by international law. AI/ML is not universally accepted in healthcare and problems encountered along the way from inception to

deployment of a diagnostic SaMD are varied and many. Changing the level of acceptance will require adherence to Good Clinical Practice. Our next objective was to compare the same two articles head-to-head with our model evaluating GCP.

Table 2: Comparisons to GCP

	Good Clinical Practice	Andreu-Perez et al. [8]	Verde et al. [9]	Our Model
	Ethics			
1	<i>Clinical trials should be conducted in accordance with ethical principles that have their origin in the Declaration of Helsinki, and that are consistent with GCP and the applicable regulatory requirement(s).</i>	Stated compliance with the Declaration of Helsinki	NR All authors are biomedical engineers or software engineers	Stated compliance with the Declaration of Helsinki
2	<i>Before a trial is initiated, foreseeable risks and inconveniences should be weighed against anticipated benefit for the individual trial subject and society. A trial should be initiated and continued only if the anticipated benefits justify the risks.</i>	Risks NR Benefits well described.	Risks NR Benefits well described	Full QMS documentation to include user and device FMEA complete before trial initiated
3	<i>The rights, safety and well-being of the trial subjects are the most important considerations and should prevail over interest of science and society.</i>	No obvious violation of rights within the article; Not fully reported with no risk analysis.	No obvious violation of rights within the article; Not fully reported with no risk analysis	Full compliance with all international and United States provisions for subject protection, subject rights, minority and underserved access/inclusion and subject safety.
4	<i>Nonclinical and clinical information supports the trial.</i>	Review of literature and related work was completed	Review of literature and related work was completed	Detailed literature reviews undertaken throughout development of model. Benchtop testing to simulate clinical trial performed until all benchmarks for trial met on benchtop.
	Protocol and Science			
5	<i>Clinical trials should be scientifically sound, and described in a clear, detailed</i>	A detailed protocol was written prior to initiating the clinical trial	No description of a clinical trial	A detailed protocol was written prior to initiating the clinical trial

	<i>protocol</i>			
	Responsibilities			
6	<i>A trial should be conducted in compliance with the protocol that has received prior institutional review board (IRB) / independent ethics committee (IEC) approval/favorable opinion.</i>	The clinical trial was approved by local IECs.	NR	The trial was approved by clinical by a national IRB and accepted by the local IRB.
7	<i>The medical care given to, and medical decisions made on behalf of subjects should always be the responsibility of a qualified physician or, when appropriate, of a qualified dentist.</i>	NR No physician listed in author list.	NR No physician listed in author list.	Clinical trial with physician PIs and Nurse Site Coordinators
8	<i>Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective task(s).</i>	RNs used to administer testing was appropriate. No other reports of qualifications.	NR All authors are biomedical engineers or software engineers	RNs and trained clinical research assistants overseen by the trial site and responsible for appropriately obtaining the samples. Designated site and sponsor physician PIs for the entire trial. Site Coordinators on site throughout the trial.
	Informed Consent			
9	<i>Freely given informed consent should be obtained from every subject prior to clinical trial participation.</i>	Full informed consent obtained on all participants.	NR	Full informed consent obtained on all participants.
	Data Quality and Integrity			
10	<i>All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation and verification</i>	GDPR compliant data collection and storage	NR	All data collected and maintained in HIPPA compliant fashion, All results reported as HL7 compliant reports in HIPPA compliant fashion. System designed to meet all medical device and SaMD ISOs, standards, good practices, and cybersecurity benchmarks
11	<i>The confidentiality of records that could</i>	Researchers ensured patient confidentiality	NR	Researchers ensured patient confidentiality during the trial.

	<i>identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).</i>	during the trial. GDPR compliant app developers designed to this end.		HIPPA/HL7 compliant app and API developers designed to this end.
	Investigational Products			
12	<i>Investigational products should be manufactured, handled and stored in accordance with applicable Good Manufacturing Practice (GMP). They should be used in accordance with the approved protocol.</i>	See Table #1 (GMLP)	See Table #1 (GMLP)	See Table #1 (GMLP)
	Quality Control / Quality Assurance			
13	<i>Systems with procedures to ensure quality of every aspect of the trial should be implemented</i>	NR	NR	Full QMS system in place and documentation complete before trial begin

NR = Not Reported

5. Discussion

The publication of GMLP that has followed the GCP for Medical Devices raises the bar for the scientific rigor of SiMD/SaMD development, the associated benchtop testing, and clinical trials. Collectively, these principles reinforce that prospective design applies not only to the design of the software but also the testing and, ultimately, the evidence-based clinical trial of the device.

Our review of 57 articles and reports describing FCV-SDS based COVID-19 diagnostic software components and systems found that only 2 of the 57 articles could meet Level of Evidence, grade of data and statistical significance of testing criteria. When these 2 articles were evaluated against the principles for GMLP and GCP, only the article by Andreu-Perez et al. demonstrated a degree of compliance with the guidelines [8].

Our own experience, albeit pre-clinical trial, found that the GMLP and GCP guidelines are well aligned with good engineering practices and good evidence-based medicine research practices. We respectfully submit the following supplements to the GMLP and GCP to include preclinical

benchtop testing principles, techniques to ensure AI/ML model consistency and guidance for overcoming data borders:

- Prior to Clinical Trial, investigational AI/ML software devices should be trained and validated using a reference dataset that is clinically relevant to the proposed clinical trial. Furthermore, the model should demonstrate successful benchtop testing using a testing dataset that is clinically relevant and with testing endpoints that are identical to prospectively designed clinical trial. This will promote adherence to two of the current GMLP principles - separation of training and test sets and providing for the highest quality of data.
- The trained and validated AI/ML models used in pre-clinical trial benchtop testing and for clinical trials should be reproducible by independent researchers. This requires that the source code for the complete model, including post training transfer learning, is made available as part of peer-reviewed published reports to facilitate replication of the benchtop testing and/or clinical trials. The use of original source code and transfer learning ensures that the replicated tests and trials are using the identical AI/ML model, making direct comparisons possible.
- Clinical trials that cross data borders and parallel AI/ML models operating in different data boundary regions should implement the same source code and transfer learning in all data regions. Additionally, if the clinical trial crosses data borders, parallel AI/ML models operating in different data boundary regions should synchronize transfer learning using federated data and swarm learning. This approach would respect applicable data boundaries, which is essential for international operations. Without these considerations, data borders will impose healthcare inequities and limit access.

6. Conclusion

Adherence to GMLP and GCP by AI/ML SaMD and AI/ML SiMD developers will also serve to improve the confidence of healthcare providers and healthcare decision makers in these new and promising technologies. The newly released GMLP guidelines provide an ideal framework for AI/ML and SiMD/SaMD developers to guide their efforts throughout the TPLC, yielding more meaningful clinical outcomes. Data scientists, clinicians, and regulatory specialists must collaborate to ensure that good practices translate into a product that meets the needs of both healthcare providers and patients. Focusing on quality and incorporating regulatory guidance throughout the TPLC will allow AI/ML developers to bring a high-quality medical to market.

GCP principles provide the guidance for clinical trials to be conducted ethically and scientifically with clear and detailed protocols. Care must be given by appropriately qualified personnel with adequate experience. Most importantly, documentation of adherence to these practices is key, with records easily accessible and retrievable for accurate reporting, verification and interpretation.

In order to bring new AI/ML technology solutions to healthcare, developers must prospectively plan to use the GMLP and GCP principles from the beginning of the TPLC.

7. References

7.1. Current Article

- [1] Coppock, H., Jones, L., Kiskin, I., & Schuller, B. (2021). COVID-19 detection from audio: seven grains of salt. *Lancet Digit Health*, 3(9), e537-e538. doi:10.1016/S2589-7500(21)00141-2
- [2] Khanzada, A., et al., *Challenges and Opportunities in Deploying COVID-19 Cough AI Systems*. J Voice, 2021.
- [3] Schuller, B.W., et al., *COVID-19 and Computer Audition: An Overview on What Speech & Sound Analysis Could Contribute in the SARS-CoV-2 Corona Crisis*. *Frontiers in Digital Health*, 2021. 3: p. 14
- [4] Topol, E. J. (2020). Is my cough COVID-19? *Lancet*, 396(10266), 1874. doi:10.1016/S0140-6736(20)32589-7
- [5] U.S. Food and Drug Administration. Good Machine Learning Practice for Medical Device Development: Guiding Principles, <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles>; 2021 (accessed 4 November 2021)
- [6] U.S. Food and Drug Administration. Good Clinical Practice 101, <https://www.fda.gov/files/medical%20devices/published/Presentation--Good-Clinical-Practice-101--An-Introduction-%28PDF-Version%29.pdf> (accessed 11 November 2021)
- [7] U.S. Food and Drug Administration. Good Clinical Practices, 62 Fed.Reg.90 (May 9, 1997) <https://www.govinfo.gov/content/pkg/FR-1997-05-09/html/97-12138.htm> (accessed 11 November 2021)
- [8] Andreu-Perez, J., Pérez-Espinosa, H., Timonet, E., Kiani, M., Giron-Perez, M. I., Benitez-Trinidad, A. B., . . . Reyes-Galaviz, O. F. (2021). A generic deep learning based cough analysis system from clinically validated samples for point-of-need covid-19 test and severity levels. *IEEE Transactions on Services Computing*.
- [9] Verde, L., et al., *Exploring the Use of Artificial Intelligence Techniques to Detect the Presence of Coronavirus Covid-19 Through Speech and Voice Analysis*. *IEEE Access*, 2021. 9: p. 65750-65757.

7.2. Bibliography of the Systematic Review of the Literature

- Adans-Dester, C. P., Bamberg, S., Bertacchi, F. P., Caulfield, B., Chappie, K., Demarchi, D., . . . Freni, M. (2020). Can mHealth technology help mitigate the effects of the COVID-19 pandemic? *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 243-248.
- Akman, A., Coppock, H., Gaskell, A., Tzirakis, P., Jones, L., & Schuller, B. W. (2021). Evaluating the COVID-19 Identification ResNet (CIdER) on the INTERSPEECH COVID-19 from Audio Challenges. *arXiv preprint arXiv:2107.14549*.
- Alkhodari, M., & Khandoker, A. (2021). Detection of COVID-19 in smartphone-based breathing recordings using CNN-BiLSTM: a pre-screening deep learning tool. *medRxiv*.

- Amoh, J., & Odame, K. (2016). Deep Neural Networks for Identifying Cough Sounds. *IEEE Trans Biomed Circuits Syst*, 10(5), 1003-1011. doi:10.1109/TBCAS.2016.2598794
- Andreu-Perez, J., Pérez-Espinosa, H., Timonet, E., Kiani, M., Giron-Perez, M. I., Benitez-Trinidad, A. B., . . . Reyes-Galaviz, O. F. (2021). A generic deep learning based cough analysis system from clinically validated samples for point-of-need covid-19 test and severity levels. *IEEE Transactions on Services Computing*.
- Bagad, P., Dalmia, A., Doshi, J., Nagrani, A., Bhamare, P., Mahale, A., . . . Panicker, R. (2020). Cough against COVID: Evidence of COVID-19 signature in cough sounds. *arXiv preprint arXiv:2009.08790*.
- Bales, C., Nabeel, M., John, C. N., Masood, U., Qureshi, H. N., Farooq, H., . . . Imran, A. (2020). *Can machine learning be used to recognize and diagnose coughs?* Paper presented at the 2020 International Conference on e-Health and Bioengineering (EHB).
- Bartl-Pokorny, K. D., Pokorny, F. B., Batliner, A., Amiriparian, S., Semertzidou, A., Eyben, F., . . . Wehler, M. (2021). The voice of COVID-19: Acoustic correlates of infection in sustained vowels. *The Journal of the Acoustical Society of America*, 149(6), 4377-4383.
- Belkacem, A. N., Ouhbi, S., Lakas, A., Benkhelifa, E., & Chen, C. (2021). End-to-End AI-Based Point-of-Care Diagnosis System for Classifying Respiratory Illnesses and Early Detection of COVID-19: A Theoretical Framework. *Front Med (Lausanne)*, 8, 585578. doi:10.3389/fmed.2021.585578
- Bhateja, V., Taqee, A., & Sharma, D. K. (2019). *Pre-processing and classification of cough sounds in noisy environment using SVM*. Paper presented at the 2019 4th International Conference on Information Systems and Computer Networks (ISCON).
- Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombat, A., Spathis, D., . . . Mascolo, C. (2020). Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. *arXiv preprint arXiv:2006.05919*.
- Burns, P. B., Rohrich, R. J., & Chung, K. C. (2011). The Levels of Evidence and Their Role in Evidence-Based Medicine. *Plastic and reconstructive surgery*, 128(1), 305-310.
- Chang, J., Cui, S., & Feng, M. (2021). DiCOVA-Net: Diagnosing COVID-19 using Acoustics based on Deep Residual Network for the DiCOVA Challenge 2021. *arXiv preprint arXiv:2107.06126*.
- Chen, X., Hu, M., & Zhai, G. (2021). Cough Detection Using Selected Informative Features from Audio Signals. *arXiv preprint arXiv:2108.03538*.
- Chetupalli, S. R., Krishnan, P., Sharma, N., Muguli, A., Kumar, R., Nanda, V., . . . Ganapathy, S. (2021). Multi-modal Point-of-Care Diagnostics for COVID-19 Based On Acoustics and Symptoms. *arXiv preprint arXiv:2106.00639*.
- Chowdhury, M. E., Ibtehaz, N., Rahman, T., Mekki, Y. M. S., Qibalwey, Y., Mahmud, S., . . . Al-Maadeed, S. A. S. (2021). QUCoughScope: An Artificially Intelligent Mobile Application to Detect Asymptomatic COVID-19 Patients using Cough and Breathing Sounds. *arXiv preprint arXiv:2103.12063*.

- Chowdhury, N. K., Kabir, M. A., & Rahman, M. (2021). An Ensemble-based Multi-Criteria Decision Making Method for COVID-19 Cough Classification. *arXiv preprint arXiv:2110.00508*.
- Cohen-McFarlane, M., Goubran, R., & Knoefel, F. (2019). *Comparison of silence removal methods for the identification of audio cough events*. Paper presented at the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
- Cohen-McFarlane, M., Goubran, R., & Knoefel, F. (2020). Novel coronavirus cough database: NoCoCoDa. *IEEE Access*, 8, 154087-154094.
- Coppock, H., Gaskell, A., Tzirakis, P., Baird, A., Jones, L., & Schuller, B. (2021). End-to-end convolutional neural network enables COVID-19 detection from breath and cough audio: a pilot study. *BMJ Innov*, 7(2), 356-362. doi:10.1136/bmjinnov-2021-000668
- Coppock, H., Jones, L., Kiskin, I., & Schuller, B. (2021). COVID-19 detection from audio: seven grains of salt. *Lancet Digit Health*, 3(9), e537-e538. doi:10.1016/S2589-7500(21)00141-2
- Deshpande, G., & Schuller, B. (2020). An overview on audio, signal, speech, & language processing for covid-19. *arXiv preprint arXiv:2005.08579*.
- Deshpande, G., & Schuller, B. W. (2020). Audio, Speech, Language, & Signal Processing for COVID-19: A Comprehensive Overview. *arXiv preprint arXiv:2011.14445*.
- Faezipour, M., & Abuzneid, A. (2020). Smartphone-based self-testing of COVID-19 using breathing sounds. *Telemedicine and e-Health*, 26(10), 1202-1205.
- Feng, K., He, F., Steinmann, J., & Demirkiran, I. (2021). *Deep-learning Based Approach to Identify Covid-19*. Paper presented at the SoutheastCon 2021.
- Granholt, A., Alhazzani, W., & Møller, M. H. (2019). Use of the GRADE approach in systematic reviews and guidelines. *British journal of anaesthesia*, 123(5), 554-559.
- Hossain, M. Z., Uddin, M. B., & Ahmed, K. A. (2021). CovidEnvelope: A Fast Automated Approach to Diagnose COVID-19 from Cough Signals. *medRxiv*.
- Hoyos-Barcelo, C., Monge-Alvarez, J., Zeeshan Shakir, M., Alcaraz-Calero, J. M., & Casaseca-de-la-Higuera, P. (2018). Efficient k-NN Implementation for Real-Time Detection of Cough Events in Smartphones. *IEEE J Biomed Health Inform*, 22(5), 1662-1671. doi:10.1109/JBHI.2017.2768162
- Imran, A., Posokhova, I., Qureshi, H. N., Masood, U., Riaz, M. S., Ali, K., . . . Nabeel, M. (2020). AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Inform Med Unlocked*, 20, 100378. doi:10.1016/j.imu.2020.100378
- Jayachitra, V. P., Nivetha, S., Nivetha, R., & Harini, R. (2021). A cognitive IoT-based framework for effective diagnosis of COVID-19 using multimodal data. *Biomed Signal Process Control*, 70, 102960. doi:10.1016/j.bspc.2021.102960

- Khazada, A., Hegde, S., Sreeram, S., Bower, G., Wang, W., Mediratta, R. P., . . . Rameau, A. (2021). Challenges and Opportunities in Deploying COVID-19 Cough AI Systems. *J Voice*. doi:10.1016/j.jvoice.2021.08.009
- Knocikova, J., Korpas, J., Vrabec, M., & Javorka, M. (2008). Wavelet analysis of voluntary cough sound in patients with respiratory diseases. *J Physiol Pharmacol*, 59(Suppl 6), 331-340.
- Laguarta, J., Hueto, F., & Subirana, B. (2020). COVID-19 artificial intelligence diagnosis using only cough recordings. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 275-281.
- Lella, K. K., & Pja, A. (2021). Automatic COVID-19 disease diagnosis using 1D convolutional neural network and augmentation with human respiratory sound based on parameters: cough, breath, and voice. *AIMS Public Health*, 8(2), 240-264. doi:10.3934/publichealth.2021019
- Liu, J.-M., You, M., Wang, Z., Li, G.-Z., Xu, X., & Qiu, Z. (2014). *Cough detection using deep neural networks*. Paper presented at the 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- Melek, M. (2021). Diagnosis of COVID-19 and non-COVID-19 patients by classifying only a single cough sound. *Neural Comput Appl*, 1-12. doi:10.1007/s00521-021-06346-3
- Melek Manshouri, N. (2021). Identifying COVID-19 by using spectral analysis of cough recordings: a distinctive classification study. *Cognitive Neurodynamics*. doi:10.1007/s11571-021-09695-w
- Miranda, I. D., Diacon, A. H., & Niesler, T. R. (2019). *A comparative study of features for acoustic cough detection using deep architectures*. Paper presented at the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
- Mohammed, E. A., Keyhani, M., Sanati-Nezhad, A., Hejazi, S. H., & Far, B. H. (2021). An ensemble learning approach to digital corona virus preliminary screening from cough sounds. *Sci Rep*, 11(1), 15404. doi:10.1038/s41598-021-95042-2
- Monge-Alvarez, J., Hoyos-Barcelo, C., Lesso, P., & Casaseca-de-la-Higuera, P. (2019). Robust Detection of Audio-Cough Events Using Local Hu Moments. *IEEE J Biomed Health Inform*, 23(1), 184-196. doi:10.1109/JBHI.2018.2800741
- Monge-Alvarez, J., Hoyos-Barcelo, C., Lesso, P., Escudero, J., Dahal, K., & Casaseca-de-la-Higuera, P. (2016). Effect of importance sampling on robust segmentation of audio-cough events in noisy environments. *Annu Int Conf IEEE Eng Med Biol Soc*, 2016, 3740-3744. doi:10.1109/EMBC.2016.7591541
- Monge-Álvarez, J., Hoyos-Barceló, C., San-José-Revuelta, L. M., & Casaseca-de-la-Higuera, P. (2018). A machine hearing system for robust cough detection based on a high-level representation of band-specific audio features. *IEEE Transactions on Biomedical Engineering*, 66(8), 2319-2330.

- Mouawad, P., Dubnov, T., & Dubnov, S. (2021). Robust Detection of COVID-19 in Cough Sounds. *SN Computer Science*, 2(1), 1-13.
- Muguli, A., Pinto, L., Sharma, N., Krishnan, P., Ghosh, P. K., Kumar, R., . . . Ramoji, S. (2021). DiCOVA Challenge: Dataset, task, and baseline system for COVID-19 diagnosis using acoustics. *arXiv preprint arXiv:2103.09148*.
- Mydukuri, R. V., Kallam, S., Patan, R., Al-Turjman, F., & Ramachandran, M. (2021). Deming least square regressed feature selection and Gaussian neuro-fuzzy multi-layered data classifier for early COVID prediction. *Expert Syst*, e12694. doi:10.1111/exsy.12694
- Nguyen, L. H., Pham, N. T., Do, V. H., Nguyen, L. T., Nguyen, T. T., Do, V. D., . . . Nguyen, N. D. (2021). Fruit-CoV: An Efficient Vision-based Framework for Speedy Detection and Diagnosis of SARS-CoV-2 Infections Through Recorded Cough Sounds. *arXiv preprint arXiv:2109.03219*.
- Orlandic, L., Teijeiro, T., & Atienza, D. (2021). The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms. *Sci Data*, 8(1), 156. doi:10.1038/s41597-021-00937-4
- Pahar, M., Klopper, M., Warren, R., & Niesler, T. (2021). COVID-19 cough classification using machine learning and global smartphone recordings. *Comput Biol Med*, 135, 104572. doi:10.1016/j.combiomed.2021.104572
- Pahar, M., Klopper, M., Warren, R., & Niesler, T. (2021). COVID-19 Detection in Cough, Breath and Speech using Deep Transfer Learning and Bottleneck Features. *arXiv preprint arXiv:2104.02477*.
- Pal, A., & Sankarasubbu, M. (2021). *Pay attention to the cough: Early diagnosis of COVID-19 using interpretable symptoms embeddings with cough sound signal processing*. Paper presented at the Proceedings of the 36th Annual ACM Symposium on Applied Computing.
- Pinkas, G., Karny, Y., Malachi, A., Barkai, G., Bachar, G., & Aharonson, V. (2020). SARS-CoV-2 detection from voice. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 268-274.
- Pizzo, D. T., Esteban, S., & Scetta, M. (2021). IATos: AI-powered pre-screening tool for COVID-19 from cough audio samples. *arXiv preprint arXiv:2104.13247*.
- Quatieri, T. F., Talkar, T., & Palmer, J. S. (2020). A framework for biomarkers of COVID-19 based on coordination of speech-production subsystems. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 203-206.
- Ritwik, K. V. S., Kalluri, S. B., & Vijayaseenan, D. (2020). COVID-19 patient detection from telephone quality speech data. *arXiv preprint arXiv:2011.04299*.
- Schuller, B. W., Batliner, A., Bergler, C., Mascolo, C., Han, J., Lefter, I., . . . Stappen, L. (2021). The INTERSPEECH 2021 Computational Paralinguistics Challenge: COVID-19 cough, COVID-19 speech, escalation & primates. *arXiv preprint arXiv:2102.13468*.

- Schuller, B. W., Coppock, H., & Gaskell, A. (2020). Detecting COVID-19 from breathing and coughing sounds using deep neural networks. *arXiv preprint arXiv:2012.14553*.
- Schuller, B. W., Schuller, D. M., Qian, K., Liu, J., Zheng, H., & Li, X. (2021). COVID-19 and Computer Audition: An Overview on What Speech & SoundAnalysis Could Contribute in theSARS-CoV-2 Corona Crisis. *Frontiers in Digital Health*, 3, 14.
- Schünemann, H. J., Oxman, A. D., Brozek, J., Glasziou, P., Jaeschke, R., Vist, G. E., . . . Montori, V. M. (2008). Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *Bmj*, 336(7653), 1106-1110.
- Shams, A. B., Raihan, M., Sarker, M., Khan, M., Uddin, M., Preo, R. B., & Monjur, O. (2021). Telehealthcare and Covid-19: A Noninvasive & Low Cost Invasive, Scalable and Multimodal Real-Time Smartphone Application for Early Diagnosis of SARS-CoV-2 Infection. *arXiv preprint arXiv:2109.07846*.
- Sharma, M., Shenoy, N., Doshi, J., Bagad, P., Dalmia, A., Bhamare, P., . . . Panicker, R. (2021). Impact of data-splits on generalization: Identifying COVID-19 from cough and context. *arXiv preprint arXiv:2106.03851*.
- Sharma, N., Krishnan, P., Kumar, R., Ramoji, S., Chetupalli, S. R., Ghosh, P. K., & Ganapathy, S. (2020). Coswara--A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis. *arXiv preprint arXiv:2005.10548*.
- Shimon, C., Shafat, G., Dangoor, I., & Ben-Shitrit, A. (2021). Artificial intelligence enabled preliminary diagnosis for COVID-19 from voice cues and questionnaires. *J Acoust Soc Am*, 149(2), 1120. doi:10.1121/10.0003434
- Tena, A., Claria, F., & Solsona, F. (2022). Automated detection of COVID-19 cough. *Biomed Signal Process Control*, 71, 103175. doi:10.1016/j.bspc.2021.103175
- Ting, D. S. W., Carin, L., Dzau, V., & Wong, T. Y. (2020). Digital technology and COVID-19. *Nature medicine*, 26(4), 459-461.
- Topol, E. J. (2020). Is my cough COVID-19? *Lancet*, 396(10266), 1874. doi:10.1016/S0140-6736(20)32589-7
- Verde, L., De Pietro, G., Ghoneim, A., Alrashoud, M., Al-Mutib, K. N., & Sannino, G. (2021). Exploring the Use of Artificial Intelligence Techniques to Detect the Presence of Coronavirus Covid-19 Through Speech and Voice Analysis. *IEEE Access*, 9, 65750-65757.
- Xue, H., & Salim, F. D. . (2021). Exploring Self-Supervised Representation Ensembles for COVID-19 Cough Classification. *arXiv preprint arXiv:2105.07566*.