Adversarially Robust Multimodal Medical Agents Integrating PPG Foundation Models for Secure Clinical Decision Support

Roben Rliver
Department of Computer Science and Engineering, University of Nevada, Reno, Reno, NV, USA.
helloruben@unr.edu

Niklas C. Hawkins
Department of Computer Science, University of Houston, Houston, TX, USA.
hawkins1982@uh.edu

Abstract

The accelerating integration of artificial intelligence into clinical workflows demands architectures that not only improve diagnostic precision but also guarantee operational security under adversarial conditions. This paper presents a comprehensive systems-level analysis of adversarially robust multimodal medical agents that incorporate photoplethysmography (PPG) foundation models for secure clinical decision support. By combining large language model reasoning with continuous, self-supervised representations of cardiovascular dynamics, these agents offer a pathway toward contextually aware, real-time patient monitoring. However, the multimodal fusion surface expands the attack vectors available to adversaries, including evasion perturbations on physiological signals, prompt injection into clinical dialogue, and data poisoning targeting pretraining pipelines. We examine the structural trade-offs inherent in designing such agents, focusing on the interaction between the PPG foundation model's inductive biases, the orchestration layer's adversarial hardening, and the governance frameworks required for clinical deployment. A layered defense architecture is proposed, integrating statistical-prior-informed robust representation learning, certified input filtering, and runtime cross-modal verification. System-level evaluation criteria are discussed in terms of clinical utility, robustness-accuracy trade-offs, latency constraints, and equity across diverse patient populations. The work further elaborates on regulatory alignment, liability distribution, and the long-term sustainability of secure medical AI ecosystems. By bridging foundational physiological modeling with adversarial machine learning, the paper contributes a blueprint for next-generation clinical agents that are both intelligent and resilient.

Keywords

multimodal medical agents, PPG foundation models, adversarial robustness, clinical decision support, secure healthcare AI, physiological foundation models, system governance.

1. Introduction

The contemporary transformation of healthcare delivery is increasingly mediated by artificial intelligence systems that promise to augment clinical reasoning, reduce cognitive load, and enable earlier detection of adverse physiological events. Large language models have demonstrated nascent capabilities in summarizing electronic health records, generating

differential diagnoses, and powering conversational agents for patient triage. In parallel, the proliferation of wearable sensors has generated a deluge of continuous physiological data, with photoplethysmography emerging as a particularly accessible modality for monitoring cardiac rhythm, blood oxygen saturation, and vascular dynamics. The confluence of these trends suggests a new class of clinical decision support system: the multimodal medical agent, capable of fusing unstructured textual reasoning with dense, temporally structured biosignals to produce holistic clinical assessments. Yet the very sophistication that makes such agents clinically attractive also renders them susceptible to a broad range of adversarial manipulations, from imperceptible signal perturbations that suppress arrhythmia detection to maliciously crafted prompts that subvert therapeutic recommendations. This paper articulates a system-oriented research agenda for building adversarially robust multimodal medical agents that integrate PPG foundation models as a secure physiological perception backbone. It moves beyond isolated model-level defenses to examine the architectural, operational, and governance dimensions that collectively determine the trustworthiness of these agents in high-stakes environments.

The central argument advanced is that robustness in clinical AI is not an additive property that can be bolted on after development; rather, it must be woven into the very fabric of the agent's multimodal integration strategy, the pretraining objectives of its physiological encoders, and the runtime orchestration logic that governs when and how recommendations are escalated to human clinicians. We constrain our discourse to system-level considerations, deliberately avoiding the mathematical formalisms that dominate the adversarial robustness literature, in order to foreground the structural trade-offs that architects and policymakers must navigate. The integration of PPG foundation models is particularly illustrative because these models encode strong prior knowledge about cardiovascular physiology, knowledge that can simultaneously serve as a robustness asset and a domain-specific attack surface. As the paper unfolds, we elaborate on the mutual dependencies between the agent's perceptual fidelity, its reasoning coherence, and the adversarial resilience of the overall clinical pipeline.

2. Background and Related Work

Clinical decision support systems have evolved from rule-based engines to data-driven models that leverage electronic health records, medical imaging, and laboratory values. The emergence of foundation models in medical imaging and clinical language processing, exemplified by domain-adapted architectures such as Med-PaLM and RadGraph, has demonstrated the potential for generalizable knowledge transfer across tasks. However, most of these efforts treat modalities in isolation, neglecting the continuous physiological streams that are increasingly available from intensive care units and consumer wearables. Photoplethysmography, an optoelectronic technique that measures blood volume changes, has historically been analyzed using handcrafted features or shallow machine learning pipelines. Recent work on PPG foundation models represents a paradigm shift, applying self-supervised pretraining to vast corpora of raw photoplethysmographic signals in order to learn universal representations that transfer to downstream tasks such as atrial fibrillation detection, blood pressure estimation, and stress quantification. The statistical-prior informed generative masking architecture described by Guo et al. [8] illustrates how inductive biases derived from cardiovascular physiology can be embedded into the model's learning objective, yielding representations that are both accurate and semantically coherent.

Adversarial machine learning in healthcare has attracted growing attention, with early work demonstrating that medical images can be perturbed to induce misdiagnosis and that clinical

text classifiers are vulnerable to synonym substitution attacks. Large language model agents introduce additional vulnerabilities, including goal hijacking, prompt injection, and the exploitation of instruction-following behaviors to generate harmful content disguised as clinical advice. Research into adversarial robustness for medical LLM agents has produced a variety of defense methodologies, ranging from safety alignment tuning to context-aware output filtering. Hu [18] provides a detailed taxonomy of security enhancement methods specifically tailored to LLM agents performing medical decision-making tasks, underscoring the necessity of adversarial training alongside architectural safeguards. Despite these advances, the literature remains largely partitioned: PPG robustness, NLP adversarial defense, and multimodal fusion security are treated as separate subfields, with minimal cross-pollination. The present work addresses this gap by considering the integrated system as the unit of analysis, emphasizing how vulnerabilities propagate across modal boundaries and how defenses can be co-designed.

3. Multimodal Medical Agent Architecture

The reference architecture we propose comprises four primary subsystems: a multimodal perception layer, a reasoning core built on a large language model, an action interface that mediates clinical communication, and an orchestration module that manages information flow and exception handling. The perception layer is responsible for ingesting and normalizing heterogeneous data streams, including free-text clinical notes, structured laboratory results, and continuous PPG waveforms. Instead of treating the PPG stream as a raw time series to be passed directly to a downstream predictor, the architecture routes it through a pre-trained PPG foundation model that transforms the signal into a compact, semantically rich embedding vector. This vector encapsulates information about heart rate variability, arrhythmic patterns, and hemodynamic stability, all inferred through the lens of the model's large-scale cardiovascular pretraining. The reasoning core receives this physiological embedding alongside textual evidence and engages in chain-of-thought deliberation, producing a clinical assessment that is both contextually grounded and physiologically informed.

Several structural trade-offs manifest at the architectural level. One pivotal choice concerns the stage at which multimodal fusion occurs. Early fusion concatenates the PPG embedding directly into the LLM's token sequence, enabling deep cross-attention between linguistic and physiological representations but dramatically increasing the prompt length and the corresponding attack surface for adversarial tokens. Late fusion preserves independent reasoning pipelines and combines likelihoods only at the decision level, which can improve robustness through redundancy but sacrifices the integrative nuance that often characterizes expert clinical reasoning. A hybrid approach, in which the PPG embedding participates in cross-modal attention within specialized adapter modules while maintaining a separate verification pathway, appears to offer a favorable balance. This design enables the agent to detect inconsistencies between the physiological signal and the clinical narrative, using such discrepancies as a trigger for fallback protocols that request human review. Latency constraints are also paramount; in tele-ICU settings, the entire perception-reasoning loop must complete within seconds to be actionable. This requirement privileges architectures that precompute PPG embeddings at the edge and stream only compact representations to the cloud-based LLM, reducing the attack surface of the transmission channel while preserving clinical responsiveness.

4. PPG Foundation Models for Physiological Sensing

Photoplethysmography offers a uniquely scalable window into cardiovascular function, yet its signals are notoriously susceptible to motion artifacts, ambient light interference, and inter-individual variability in skin pigmentation and perfusion. Traditional PPG analytics rely on peak detection algorithms that degrade sharply under non-ideal conditions. Foundation models pre-trained on diverse, multi-site datasets address these challenges by learning to disentangle biological signal from environmental noise in a self-supervised manner. The generative masking approach exemplified by the SIGMA-PPG framework [8] endows the model with statistical priors that reflect physiological regularities, such as the quasi-periodic nature of cardiac cycles and the limited bandwidth of hemodynamic fluctuations. By reconstructing intentionally masked segments of the PPG waveform, the model internalizes a robust generative model of cardiovascular dynamics that can be fine-tuned for a wide array of clinical tasks with limited labeled data.

Integrating such a foundation model into a multimodal agent confers several architectural advantages. The resulting embeddings are inherently smooth with respect to small input perturbations, a property that arises from the model's training objective and can be further enhanced through randomized smoothing layers. This smoothness provides a degree of inherent adversarial robustness, making it more difficult for an attacker to craft a perturbation that simultaneously evades detection by the foundation model and remains physiologically plausible. Nonetheless, a determined adversary can exploit domain-specific artifacts, such as simulating motion corruption in a manner that masks a true arrhythmic episode without triggering outlier detection. The agent must therefore couple the PPG foundation model with contextual plausibility checks that monitor the coherence of the physiological embedding with other sensor streams, such as electrocardiogram data when available, or with the expected hemodynamic response to documented clinical interventions. The data governance dimensions are equally critical; PPG data often originates from consumer wearables that reside outside the hospital's administrative domain, raising questions about provenance, consent, and the potential for poisoning attacks during federated fine-tuning cycles.

5. Adversarial Threat Landscape in Clinical AI

The security of a multimodal clinical agent must be evaluated against a comprehensive threat model that accounts for the motivations and capabilities of adversaries ranging from financially motivated fraudsters to state-sponsored actors targeting critical healthcare infrastructure. Data poisoning is a persistent concern for both the PPG foundation model and the LLM reasoning core. If an adversary controls a subset of training PPG recordings, they can implant backdoors that cause the model to systematically fail to detect a dangerous arrhythmia when a specific trigger pattern, such as a subtle high-frequency oscillation, is present in the signal. In the textual domain, poisoning can embed trigger phrases that later cause the LLM to deviate from clinical guidelines when queried. Evasion attacks at inference time pose a distinct category of threat: an adversary with access to the patient's monitoring device, perhaps through a compromised smartphone application, could inject adversarial perturbations into the PPG signal that are imperceptible to the human eye yet sufficient to flip the agent's diagnostic output. The transferability of such perturbations across different foundation model architectures compounds the risk, as an attack developed on a public model could potentially generalize to a proprietary deployment.

The multimodal fusion layer introduces novel attack vectors that do not exist in unimodal systems. An adversary can craft a scenario where the textual clinical note and the perturbed PPG signal are mutually consistent in a misleading direction, exploiting the agent's learned

tendency to upweight evidence that aligns across modalities. For example, a false narrative of stable cardiovascular status in the clinical note, combined with a perturbed PPG signal that suppresses the signature of decompensation, creates a self-reinforcing loop that is more potent than either manipulation alone. The agent's chain-of-thought reasoning, while valuable for interpretability, can be subverted by adversarially injected reasoning traces that lead the model toward a predetermined conclusion. Furthermore, membership inference attacks on the PPG foundation model's training data raise privacy concerns, as adversaries could determine whether a specific individual's physiological recording was included in the pretraining corpus, potentially revealing sensitive health information. The security enhancement methods surveyed by Hu [18] for LLM agents, including adversarial fine-tuning, safety guardrails, and input sanitization, provide a partial remedy for the textual strand of these threats, but they must be extended and adapted to account for the continuous, physiologically grounded nature of the PPG modality.

6. Integration and Robustness Enhancement Strategies

A layered defense-in-depth strategy is essential for achieving meaningful adversarial robustness in multimodal clinical agents. The innermost layer resides within the PPG foundation model itself, where adversarial training can be applied by exposing the model during pretraining or fine-tuning to perturbations generated by projected gradient descent attacks constrained to respect physiological limits. The use of statistical priors, as in the SIGMA-PPG architecture, provides an additional robustness lever, because the model's reconstructive objective bounds the space of plausible signals and renders many unnatural perturbations non-viable. Building upon this, a certified robustness wrapper can employ randomized smoothing to provide probabilistic guarantees that the embedding remains stable within a specified Lp-ball around the input, a property that is particularly valuable in regulated medical device contexts where failure mode documentation is mandatory.

The intermediate layer operates at the level of the agent's orchestration module. Here, cross-modal consistency checks serve as an anomaly detection mechanism. Before a clinical recommendation is escalated, the agent computes the expected PPG embedding distribution conditioned on the clinical narrative and compares it to the observed embedding, flagging deviations that exceed a dynamically tuned threshold. When the agent is deployed in settings with redundant sensor modalities, such as simultaneous PPG and single-lead ECG, the orchestrator can enforce stricter cross-signal verification, recognizing that the cost of additional computation is justified by the elevated security posture. For the textual reasoning pathway, the measures described in the literature on adversarially robust LLM agents are fully applicable: input detoxification through paraphrase, safety-specific fine-tuning with adversarial demonstrations, and retrieval-augmented generation that anchors outputs to authoritative medical knowledge sources all constrain the space of harmful outputs. Hu [18] details numerous such methods and evaluates their effectiveness against both white-box and black-box attack regimes, showing that a combination of adversarial training and runtime constraint enforcement significantly reduces the success rate of medical prompt injection attacks.

The outermost layer encompasses system-level monitoring and human-in-the-loop fallback protocols. A dedicated security observer module continuously profiles the statistical properties of incoming PPG streams and linguistic inputs, computing epistemic uncertainty estimates and detecting distributional shifts that may indicate an ongoing attack. Automated incident response workflows can dynamically escalate suspicious cases to a human clinical

review queue, temporarily degrade the agent's autonomy level, or invoke a fallback analytics pipeline that relies on simpler, interpretable features with a narrower attack surface. The trade-off between security and clinical utility must be carefully calibrated, as overly aggressive anomaly detection thresholds lead to alert fatigue and can erode clinician trust. Deployment architectures that co-locate the PPG foundation model and the early-stage anomaly detection on an edge device reduce the network-level attack surface and ensure that basic cardiovascular safety monitoring continues even if the cloud connection is disrupted.

7. System-Level Evaluation and Trade-offs

Evaluating the integrated system requires moving beyond component-level metrics such as adversarial accuracy under a specific perturbation budget to embrace multidimensional, clinically grounded assessments. Clinical utility can be quantified through decision curve analysis that measures net benefit across a range of risk thresholds, accounting for both appropriate interventions and false alarms. Adversarial robustness should be evaluated under realistic threat models that simulate actual physiological attack capabilities, including constraints on perturbation amplitude, frequency spectrum, and temporal coherence. A system that is robust to unbounded L-infinity perturbations but fails against motion artifact simulation offers false reassurance. The inevitable trade-off between clean-data performance and adversarial robustness merits explicit characterization: adversarial training of the PPG foundation model may slightly reduce its sensitivity to subtle but genuine hemodynamic variations, while aggressive input filtering in the LLM pathway can degrade conversational fluency and increase reasoning latency.

Deployment considerations further shape the evaluation landscape. In a tele-ICU context, a multimodal agent might process a continuous PPG feed from a wearable patch alongside hourly clinical notes and medication administration records. The system's latency budget, typically on the order of a few seconds for critical arrhythmia alerts, requires that the PPG embedding computation, cross-modal verification, and LLM reasoning all execute within a tightly orchestrated pipeline. Energy efficiency becomes a constraint for the edge component, necessitating model quantization or knowledge distillation that may inadvertently alter the robustness profile. The regulatory pathway for such a system would likely fall under FDA guidance for Software as a Medical Device, requiring comprehensive validation of both the individual component models and their integrated behavior under foreseeable misuse scenarios. Audit log design that preserves forensic evidence of adversarial perturbations without compromising patient privacy is a non-trivial systems engineering challenge that intersects with cybersecurity logging standards.

Fairness across demographic groups introduces another axis of evaluation. PPG signal quality is known to vary with skin melanin content, peripheral perfusion, and sensor placement, raising the prospect that the foundation model's representation quality, and consequently the agent's diagnostic accuracy, systematically degrades for patient subgroups that are underrepresented in the pretraining corpus. Adversarial robustness can amplify these disparities if defenses are calibrated predominantly on data from well-represented populations, creating a situation where minority patients are simultaneously more likely to experience misdiagnosis and more vulnerable to targeted attacks. Mitigating this requires stratified adversarial evaluation, diverse data collection initiatives, and algorithmic fairness constraints that are incorporated into the robustness optimization loop directly, rather than treated as a post-hoc remediation step.

8. Governance, Fairness, and Policy Implications

The deployment of adversarially robust multimodal medical agents raises profound governance questions that extend well beyond technical verification. The allocation of liability when an adversarial attack succeeds in a clinical setting remains legally unsettled. If a perturbation on a consumer wearable causes the agent to miss a fatal arrhythmia, liability could theoretically attach to the device manufacturer, the AI system developer, the healthcare institution, or the clinician who relied on the agent's output. Clear regulatory frameworks are needed that apportion responsibility based on the foreseeability of the attack vector, the degree of reasonable care exercised in system design, and the transparency of the operational boundaries communicated to users. The European Union's AI Act categorizes many medical AI applications as high-risk, imposing requirements for robustness, accuracy, and cybersecurity that are directly relevant to the systems discussed herein. However, the Act's provisions were not drafted with adversarial machine learning specifically in mind, and implementation guidance will need to evolve to address prompt injection, signal-space attacks, and the complex failure modes of multimodal fusion.

Transparency mechanisms must be designed to serve both clinical safety and regulatory oversight. Explainability techniques that trace the agent's recommendation to specific physiological features in the PPG embedding and to specific evidentiary passages in the clinical text can enable clinicians to detect inconsistencies that may signal adversarial interference. At the same time, excessive transparency can arm adversaries with information about the system's decision boundaries, enabling more potent white-box attacks. Governance structures must therefore balance openness with the imperative of security-through-obscurity where necessary, perhaps through tiered access to model internals based on authenticated roles. Institutional review boards and data safety monitoring committees will need to develop competencies in adversarial AI risk assessment, moving beyond their traditional focus on privacy and informed consent to evaluating the cybersecurity posture of integrated clinical AI systems.

Sustainability is a further governance consideration. The computing resources required for adversarial training and runtime robustness verification of large multimodal models carry a significant carbon footprint and may exacerbate healthcare cost disparities if only well-resourced institutions can deploy the most secure systems. A commitment to federated learning and model compression can mitigate these concerns, but federated learning itself introduces new attack surfaces that must be governed through secure aggregation protocols and differential privacy guarantees. International collaboration on open, adversarially hardened benchmarks for multimodal clinical AI would accelerate progress while fostering a shared understanding of acceptable risk thresholds.

9. Conclusion

This paper has presented a system-level analysis of adversarially robust multimodal medical agents that integrate PPG foundation models, advancing the thesis that clinical AI security requires a holistic architecture that spans physiological signal processing, language reasoning, cross-modal verification, and human-in-the-loop oversight. By embedding physiological priors into the foundation model's inductive biases, layering adversarial defenses across the perception and reasoning pipelines, and architecting the deployment to accommodate edge-cloud trade-offs, it is possible to construct agents that are both clinically useful and meaningfully resilient to determined adversaries. The integration of statistical-prior informed generative masking architectures for PPG and the application of security enhancement methods for LLM agents represent mutually reinforcing pillars of such a system. The

challenges that remain are substantial, spanning fairness across populations, regulatory adaptation, liability allocation, and the continuous evolution of the threat landscape. We encourage the research community to pursue interdisciplinary collaborations that link signal processing, machine learning security, human-computer interaction, and health policy, ensuring that the clinical AI systems of the next decade are designed from the ground up to earn and maintain the trust of patients and clinicians alike.

References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems (pp. 5998–6008).

2. Finlayson, S. G., Bowers, J. D., Ito, J., Zittrain, J. L., Beam, A. L., & Kohane, I. S. (2019). Adversarial attacks on medical machine learning. Science, 363(6433), 1287–1289.

3. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

4. Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H. W., Scales, N., Tanwani, A., Cole-Lewis, H., Pfohl, S., Payne, P., Seneviratne, M., Gamble, P., Kelly, C., Babiker, A., Schärli, N., Chowdhery, A., Mansfield, P., Demner-Fushman, D., … Natarajan, V. (2023). Large language models encode clinical knowledge. Nature, 620(7972), 172–180.

5. Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. Nature Medicine, 28(1), 31–38.

6. Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., & Swami, A. (2017). Practical black-box attacks against machine learning. In Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security (pp. 506–519).

7. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., & Vladu, A. (2018). Towards deep learning models resistant to adversarial attacks. In International Conference on Learning Representations.

8. Guo, Z., Chen, T., Jiao, Y., Pan, Y., Hu, X., & Ferrario, M. (2026). SIGMA-PPG: Statistical-prior Informed Generative Masking Architecture for PPG Foundation Model. arXiv preprint arXiv:2601.21031.

9. Cohen, J. M., Rosenfeld, E., & Kolter, J. Z. (2019). Certified adversarial robustness via randomized smoothing. In Proceedings of the 36th International Conference on Machine Learning (pp. 1310–1320).

10. Carlini, N., & Wagner, D. (2017). Towards evaluating the robustness of neural networks. In 2017 IEEE Symposium on Security and Privacy (pp. 39–57).

11. Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. In International Conference on Learning Representations.

12. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. In Advances in Neural Information Processing Systems.

13. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., … Liang, P. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.

14. Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). Emergent abilities of large language models. Transactions on Machine Learning Research.

15. Moor, M., Banerjee, O., Abad, Z. S. H., Krumholz, H. M., Leskovec, J., Topol, E. J., & Rajpurkar, P. (2023). Foundation models for generalist medical artificial intelligence. Nature, 616(7956), 259–265.

16. Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2019). ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. In International Conference on Learning Representations.

17. Kurakin, A., Goodfellow, I. J., & Bengio, S. (2017). Adversarial examples in the physical world. In International Conference on Learning Representations Workshop.

18. Hu, S. (2026). Research on Security Enhancement Methods for Adversarial Robust Large Language Model Intelligent Agents for Medical Decision-Making Tasks. arXiv preprint arXiv:2605.08257.

19. Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. In International Conference on Learning Representations.

20. Liang, P., Bommasani, R., Lee, T., Tsipras, D., Soylu, D., Yasunaga, M., Zhang, Y., Narayanan, D., Wu, Y., Kumar, A., Newman, B., Yuan, B., Yan, B., Zhang, C., Cosgrove, C., Manning, C. D., Ré, C., Acosta-Navas, D., Hudson, D. A., … Koreeda, Y. (2023). Holistic evaluation of language models. Transactions on Machine Learning Research.

21. Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. Current Cardiology Reviews, 8(1), 14–25.

22. Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement, 28(3), R1–R39.

23. Biswas, D., Everson, L., Liu, M., Panwar, M., Verhoef, B. E., Patki, S., Kim, C. H., Acharyya, A., Van Hoof, C., Konijnenburg, M., & Van Helleputte, N. (2019). CorNET: Deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment. IEEE Transactions on Biomedical Circuits and Systems, 13(2), 282–291.

24. Pereira, T., Tran, N., Gadhoumi, K., Pelter, M. M., Do, D. H., Lee, R. J., Colorado, R., Meisel, K., & Hu, X. (2020). Photoplethysmography based atrial fibrillation detection: A review. NPJ Digital Medicine, 3(1), 3.

25. Kumar, A., Agarwal, S., Garg, A., & Sharma, T. (2023). Adversarial attacks and defenses in clinical NLP: A systematic review. Journal of Biomedical Informatics, 140, 104316.