Advances at the intersection of engineering, physics, and medicine have shaped some of the most meaningful improvements in global health. In recognition of how these innovations influence the quality of care and the lives of patients, I have committed myself to developing the skills needed to contribute to this ongoing process. Growing up in a medically underserved region of the world, I have personally observed the persistent gap between what is technically possible and what is clinically accessible. Through my research experiences in artificial intelligence applications in medicine, in areas such as medical imaging and multimodal clinical data integration, I have developed a commitment to creating solutions that are clinically effective and broadly implementable. I aspire to be part of a generation of scientists and engineers who bridge the existing gaps, by developing technologies that address real-world healthcare challenges regardless of geographic or socioeconomic barriers. I am therefore applying to the Harvard-MIT Health Sciences and Technology PhD program in Medical Engineering and Medical Physics (MEMP) because of its unique emphasis on translational research in healthcare. I believe this program offers the best environment to continue to learn and contribute in such translational research, given its diverse faculty in patient-centered clinical research, and computational methods.

My current research interests lie at the intersection of artificial intelligence, biomedical data science, and computational medicine, with works in multimodal medical imaging analysis, and clinical data integration. My works explore three complementary directions:

(A) Foundational algorithmic methods for building trustworthy automated machine learning systems such as incorporating uncertainty quantification to flag low-confidence predictions.

(B) Representation learning across diverse biomedical data modalities including ultrasound, magnetic resonance images, and computed tomography for predictive and diagnostic modeling.

(C) Ultimately working towards accessible and reliable methods for clinical applications.

The challenges I aim to address center on the reliability, interpretability, and equitable deployment of computational methods in clinical practice. Many current systems may struggle with generalization across diverse patient populations, lack mechanisms for uncertainty awareness and remain inaccessible in settings with limited computational or clinical resources. I further want to advance methods that go beyond prediction, developing systems capable of supporting intervention-based clinical decision making through causal inference, rather than simply identifying patterns. Given my commitment to clinically grounded computation and my experience working in resource-limited contexts, I believe the HST MEMP program is a natural fit for the next steps in my training.

I am especially excited about the Medical Physics component of the program, with courses like Imaging Biophysics and Clinical Applications, and Data Acquisition & Image Reconstruction in MRI. Recognizing the importance of a good physics foundation, I have independently taken additional coursework in MRI physics to better prepare myself for translational imaging research. In my view, the computational and quantitative physics training, alongside clinical course work are indispensable for developing meaningful translational tools. Furthermore, the **Neuroimaging Training Program** offers hands-on training in advanced imaging techniques that directly supports my current research interest in neuroimaging analysis, preparing me for future work in developing physics-informed computational methods across diverse neuroimaging platforms for clinical applications.

My overarching goal is to bridge the technical and translational dimensions of artificial intelligence in medicine, developing systems that are computationally efficient, clinically meaningful, and deployable in real-world environments. In the short term, I aim to advance my expertise in biomedical data science with emphasis on model generalization, interpretability, and equitable performance across populations. In the long term, I aspire to lead research that drives data-informed healthcare innovations, translating computational and artificial intelligence innovations into deployable healthcare tools that improve diagnostic accuracy and expand global access to quality care. Beyond research, I am equally committed to mentorship and capacity building, contributing to the development of institutions in underserved communities capable of achieving globally recognized scientific impact.

My research experience spans roles at the Responsible Artificial Intelligence Lab (RAIL-KNUST), and at the Kumasi Center for Collaborative Research in Tropical Medicine (KCCR) as a research assistant and engineer, where I have worked in developing AI-driven solutions for real-world healthcare challenges. My work in these roles has been recognized at national and international venues, including the annual International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). At RAIL, I developed a good foundation in computational methods and experimental design. I worked on projects that required collecting biomedical sensor data with microcontrollers for low-resource sleep apnea monitoring, as well as exploring generative modeling approaches to improve pretraining in settings where annotated datasets are not readily available. At KCCR, work I am particularly proud of is a project on a context-specific LLM application designed to assist expectant mothers with gestational diabetes in disease management, further discussed in the next section.

My research has been guided by a consistent question: how can we build computational methods that are clinically meaningful and deployable where they are needed most? The three projects below illustrate how I have approached this balance, working toward solutions that could realistically translate to practice:

1. Non-Invasive Treatment Response Assessment and the Challenge of Generalization One of my most technically instructive projects involved the development of an end-to-end framework for breast tumor segmentation and pathological complete response (pCR) assessment in dynamic contrast-enhanced MRI [1]. This work was motivated by the increasing prevalence of breast cancers worldwide, projected to disproportionately affect low- and middle-countries in the coming years. pCR assessment is performed to evaluate a patient's response to treatment, pre-surgery. It is usually a highly invasive biopsying proceduce. I wanted to explore the possibility of performing pCR at the image-level, using extracted radiomic tumor features from the segmentation step, providing a less invasive and broadly accessible alternative. There was a substantial amount of variability in patient subgroups, which has informed my current interests in integrating demographic priors into the learning process.

2. State-of-the-Art Segmentation in Low-Resource Contexts Recently, I supervised a team that achieved 1st Place in the Brain Tumor Segmentation Challenge (BraTS-SSA) at MICCAI 2025 [2]. The challenge was characterized by a small training dataset, variable scanning quality, and significant variability in tumor presentation. We developed a segmentation-aware data augmentation pipeline combined with deep ensembling. Segmentation-aware augmentations involved performing rigorous elastic deformation transforms that are masked only within the tumor regions, creating anatomically plausible augmented samples. Through this work, I learned how context-aware methodological choices can lead to state-of-the-art performances despite limited and variable data.

3. Domain-Specific LLM Adaptation for Disease Management Expectant mothers with gestational diabetes often lack consistent access to reliable health information, particularly in rural areas where specialist consultations are infrequent. At KCCR, I led the research and engineering efforts in addressing this problem, developing a context-specific LLM application for gestational diabetes management. This required solving several interconnected problems: experimenting with vector-based and graph-based retrieval-augmented generation to ground responses in verified clinical information; fine-tuning and prompt engineering for evidence-based guidance across English and local languages; and designing a system architecture for deployment via WhatsApp to maximize accessibility. Working with clinicians, we iteratively refined and piloted the application in a controlled setting. What made this project most rewarding was witnessing the direct clinical utility, expanding access by providing evidence-based guidance to patients in their own language and on a platform they already use.

In Service to Community, I have had the honor of contributing in roles that allow me to give back and learn from those around me. Most recently, I have been involved in the Sprint AI Training Program for Medical Imaging Knowledge Translation (SPARK Academy) [1]. It is an initiative dedicated to advancing research capacity in medical imaging and deep learning especially in Africa, where I help coordinate the program and develop instructional materials for up and coming researchers and clinicians. A noteworthy outcome of this was supervising the Ghanaian team to winning first place in the Brain Tumor Segmentation Challenge (BraTS-SSA) at MICCAI 2025, marking the first time an African team has achieved this distinction [2]. My community engagements also include facilitating *SheCodes*, a girl-child empowerment program in STEM, where I teach students from the high-school to university levels in Python programming, microcontrollers, machine learning, and data science. In the open-source community, I have contributed to the TorchIO library for medical image processing, extended the nnU-Net framework to support additional use cases such as uncertainty modeling, and I am currently developing a Python library that simplifies access to medical image segmentation datasets for learning and benchmarking purposes [2]. In all these efforts, my goal is consistent: to help build communities and tools that make advanced computational methods and research more accessible, equitable, and impactful.

These experiences collectively inform my decision to pursue graduate study with the HST MEMP Program, demonstrating my preparedness through independent research efforts [1], and effective teamwork and leadership evident in collaborative projects [2] and coordination at the SPARK Training Program. I am inspired by the works of **Dr. Polina Golland**, who I had the pleasure of meeting at MICCAI, for her works in computer vision applications in medicine; **Dr. Marzyeh Ghassemi** for her works in designing fair deep learning algorithms for healthcare; **Dr. Laura Lewis** and **Dr. Lauren O'Donnell** for their works in brain imaging analysis which aligns with my interest in neuroimaging; and **Dr. Peter Szolovits** for his pioneering works in biomedical informatics at the Clinical Decision Making Group. I look forward to the opportunity of contributing my unique perspective to the HST community, and learning from its culture of collaboration, where clinicians and scientists work together to transform research into tangible improvements in healthcare.

References

[1] **Toufiq Musah**. Large kernel mednext for breast tumor segmentation and self-normalizing network for pcr classification in magnetic resonance images. In *Artificial Intelligence and Imaging for Diagnostic and Treatment Challenges in Breast Care*, pages 72–80. Springer Nature Switzerland, 2026.

[2] Claudia Takyi Ankomah, Livingstone Eli Ayivor, Ireneaus Nyame, Leslie Wambo, Patrick Yeboah Bonsu, Aondona Moses Iorumbur, Raymond Confidence, and **Toufiq Musah**. How we won brats-ssa 2025: Brain tumor segmentation in the sub-saharan african population using segmentation-aware data augmentation and model ensembling. *arXiv preprint arXiv:2510.03568*, 2025.

[1] https://event.fourwaves.com/spark/pages
[2] https://github.com/toufiqmusah/MedSegMNIST