Unlocking the Potential of Artificial Intelligence in Precision Oncology and Theranostics: Feasibility, Challenges, and Future Directions

Lalith Kumar Shiyam Sundar, PhD[1]

[1] Quantitative Imaging and Medical Physics Team, Medical University of Vienna, Vienna, Austria

Executive Summary

The International Centers for Precision Oncology (ICPO) Foundation is advancing a strategic initiative to harness artificial intelligence (AI) in precision oncology and theranostics. By leveraging the power of AI and big data, ICPO aims to enhance patient management in oncology, focusing on improving therapeutic outcomes and overall survival predictions. This approach aligns with ICPO's commitment to driving innovation in cancer care through cutting-edge technologies.

To support this vision, ICPO plans to fund innovative AI research projects through a structured research call. The call emphasizes three critical pillars: standardizing data curation, applying AI algorithms to predict outcomes, and validating AI-enhanced strategies in clinical theranostics settings. These pillars reflect the foundation's dedication to addressing key areas where AI can deliver transformative benefits in oncology.

The bespoke white paper İdentifies the challenges associated with implementing AI in precision oncology and outlines the essential considerations required for successful research and application. The findings aim to provide clarity on the potential of AI to reshape oncology practice and guide the development of impactful, data-driven solutions.

Introduction

Precision oncology stands at the frontier of modern medicine, offering unprecedented opportunities for tailored cancer diagnosis and treatment. At the heart of this transformative field lies the potential of artificial intelligence (AI) to analyze complex datasets, uncover patterns invisible to human cognition, and improve patient outcomes through predictive and personalized interventions. However, while the promise of AI in healthcare is immense, its success is fundamentally dependent on access to large, well-structured datasets.

In commercial domains, AI thrives due to the availability of vast, standardized datasets, allowing models to be trained at scale with remarkable capabilities (e.g ChatGPT). In contrast, healthcare AI faces significant barriers, particularly in oncology and theranostics. Although hospitals house immense volumes of data, these resources remain fragmented, massively unstructured and inaccessible due to strict privacy regulations and a lack of interoperability between systems. Open-source initiatives have sought to address these limitations, offering datasets that fuel AI development in fields such as

radiology. However, in theranostics and nuclear medicine, the availability of open-source data remains negligible, significantly impeding progress in AI applications for these critical domains.

Recognizing that structured data is the foundation of successful AI integration, the International Centers for Precision Oncology (ICPO) Foundation has strategically prioritized data curation in its mission to advance precision oncology and theranostics. As part of its broader initiative to harness the power of AI, ICPO has proposed a research call divided into three stages. The first stage focuses on curating large datasets in a standardized and structured manner, establishing a reliable foundation for AI algorithms. The second stage involves applying these AI models to predict clinical outcomes, emphasizing the potential for AI to enhance patient management. The final stage addresses the critical need for clinical validation, ensuring that AI-driven insights translate into tangible benefits in theranostics practice.

By dedicating a separate stage of its research call specifically to data curation, ICPO acknowledges the strategic importance of addressing the data gap that has long hindered progress in AI for molecular medicine based oncology. This white paper evaluates the feasibility of such an ambitious undertaking, identifying the challenges inherent in curating and utilizing data at scale, developing predictive AI models, and validating their clinical application. Through this exploration, it aims to provide a clear roadmap for advancing AI in precision oncology and theranostics, ensuring that the promising potential of these technologies is fully realized.

Data Curation: Prior Efforts, Challenges, and Pathways to Innovation

Over the years, several initiatives have attempted to address the critical need for accessible and standardized medical imaging datasets, driven mostly by academic institutions and non-profit organizations. These efforts have been invaluable for fields like radiology, where imaging data plays a central role in advancing diagnostics and AI applications. However, the same cannot be said for theranostics and nuclear medicine, where such open-access datasets remain disproportionately underdeveloped.

Projects like The Cancer Imaging Archive (TCIA), funded by the National Cancer Institute, have made substantial contributions by offering de-identified datasets, including CT, MRI, and PET scans, to fuel cancer research and algorithm development[1]. Similarly, the Open Access Series of Imaging Studies (OASIS) has provided critical neuroimaging data to support studies in cognitive aging and related disorders, creating opportunities for the development of AI solutions in neuroscience[2].

Other initiatives, such as the Medical Imaging and Data Resource Center (MIDRC), focus on large-scale data pooling and curation, aiming to provide a high-quality, standardized repository for advancing medical AI research[3]. Platforms like MedPix [4]and OpenNeuro [5] offer targeted datasets for education and algorithm development, while Open Data MICCAI provides resources for machine learning competitions. These efforts, while admirable, predominantly serve radiology and related domains. Theranostics, by contrast, is left underserved, with open-access nuclear medicine datasets still few and far between. Historically, nuclear medicine has not been a field known for sharing data widely. While radiology has made tremendous strides, with publicly available datasets now exceeding one million images, nuclear medicine has struggled to keep pace, with open-access datasets barely crossing the 2,000-image mark.

This disparity highlights a critical gap in the adoption of AI for theranostics and nuclear medicine, where the scarcity of large, accessible datasets limits progress.

To bridge this gap, we can take inspiration from our radiology colleagues, whose success in creating and sharing data repositories can serve as a blueprint for nuclear medicine. Adopting similar strategies—tailored to the unique requirements of theranostics—could catalyze the development of robust datasets, enabling the field to fully embrace the potential of AI.

Curating such datasets, however, is not a trivial task. It requires a well-structured and collaborative workflow, especially when working across institutions or within an established network [6]. The process typically involves several key steps:

1. **Initial Engagement**: The process begins with discussions among stakeholders—researchers, clinicians, and institutional representatives—to align on goals and expectations for the dataset.
2. **Regulatory Approvals**: Securing data transfer agreements and institutional review board (IRB) approvals is a critical next step. These ensure compliance with privacy regulations and ethical standards, particularly when working with sensitive patient data.
3. **Data Curation**: Once approvals are secured, the focus shifts to the collection and organization of data. This involves ensuring that the data is well-annotated, standardized, and formatted for compatibility with AI tools.
4. **Data Anonymization**: Protecting patient privacy is paramount. Identifiable information must be thoroughly anonymized while preserving the integrity of the data for research purposes.
5. **Data Transfer**: After curation and anonymization, the dataset is securely transferred to a centralized repository where it can be accessed by authorized researchers.

Curating healthcare datasets presents significant challenges, particularly during the regulatory approval and data curation phases. Securing data transfer agreements and institutional review board (IRB) approvals often involves extensive manual processes, leading to delays. Similarly, data curation requires meticulous effort to ensure proper annotation, standardization, and compatibility with AI frameworks. These hurdles can impede the timely development of datasets essential for innovation.

To overcome these challenges, embracing digital transformation is crucial. Implementing pre-designed IRB templates can streamline the regulatory approval process by providing institutions with ready-to-use, compliant frameworks, thereby reducing administrative burdens and expediting approvals.

In the realm of data curation, integrating generative AI offers transformative potential. Generative AI tools can automate much of the annotation and standardization work, decreasing reliance on manual labor while maintaining data quality. Recognizing this opportunity, the EIC Pathfinder Challenge 2025 has identified GenAI-powered multimodal data curation and integration as a key investment area, allocating substantial resources to support its development[7]. In this context, ICPO's funding call could serve as a valuable pilot initiative, providing researchers with preliminary data and frameworks necessary to secure EIC funding and expand their efforts.

Furthermore, adopting platforms like Gen3 can provide a standardized infrastructure for data sharing and visualization[8]. Gen3 is an open-source platform designed to build data commons and ecosystems,

offering modular services for data management, analysis, and sharing. Its framework services expose powerful APIs that facilitate interoperability within a healthy data ecosystem, enabling the development of new tools for sharing and analyzing data across resources in the cloud. By utilizing Gen3, institutions can ensure that curated data is not only of high quality but also accessible and interoperable, fostering collaboration and accelerating research in theranostics and nuclear medicine.

By implementing these strategies, the path to innovation becomes clearer. Digital transformation, combined with targeted funding and scalable platforms like Gen3, can help overcome traditional data curation hurdles, unlocking the full potential of AI for theranostics.

Applying AI to Structured Datasets: Lessons from Industry and Opportunities for Theranostics

The remarkable success of commercial AI can largely be attributed to the seamless collaboration between industry and academia in open-sourcing tools and frameworks (e.g. PyTorch[9]). From the outset, AI models and libraries were shared transparently, allowing developers and researchers to build on one another's work. This openness accelerated advancements, reduced duplication of effort, and lowered the barriers to entry for those wishing to innovate. As a result, the commercial AI space flourished, setting a powerful precedent for other domains.

Healthcare, however, has lagged behind in adopting a similarly open approach. Until recently, most healthcare AI tools were siloed, limiting collaboration and slowing progress. Encouragingly, a few initiatives have started to reverse this trend. In radiology, for example, tools like nnUNet[10], TotalSegmentator[11], and NVIDIA's MONAI [12]have set new standards for transparency and collaboration by making their codebases and models publicly available. These open frameworks have not only advanced radiological AI but have also attracted substantial investment from major technology companies, demonstrating the value of shared infrastructure in driving innovation.

By contrast, nuclear medicine and theranostics remain underrepresented in this movement toward openness. There are some promising exceptions, such as the Enhance.PET initiative[13], where both models and source code are made fully open, but such efforts are few and far between. This disparity underscores the need for theranostics to adopt similar practices and embrace the lessons learned from both radiology and the broader commercial AI space.

Sharing AI tools and models in an open and transparent manner is not only feasible but also cost-effective. Open-sourcing accelerates the development and adoption of AI by enabling researchers to validate, refine, and extend existing solutions. For theranostics, this approach could be transformative, fostering collaboration and ensuring that advancements benefit the field as a whole rather than remaining siloed within individual organizations.

To drive this shift, ICPO's funding call should mandate the open-sourcing of AI solutions developed through its support. By requiring researchers to share their models and code, ICPO can set a standard for transparency in theranostics and nuclear medicine, encouraging a culture of collaboration akin to that which has propelled radiology forward. This directive would not only benefit current projects but also establish a foundation for future innovation, ensuring that the tools and insights generated today remain accessible and impactful in the years to come.

Clinical Deployment of AI: Evaluating Real-World Utility

The true measure of an AI algorithm's success lies not only in its development but also in its deployment within clinical settings, where its real-world utility can be rigorously evaluated. This final stage of the ICPO call is dedicated to ensuring that AI models move beyond theoretical potential and demonstrate tangible benefits in theranostics and nuclear medicine. With the right tools and strategies, deploying AI algorithms in clinical environments is both feasible and impactful.

Recent advances in open-source solutions have made this process significantly more streamlined. Platforms like MONAI Deploy offer an end-to-end framework for transitioning AI models from research into clinical use[14]. MONAI Deploy simplifies the development, packaging, and deployment of AI applications through its MONAI Application Package (MAP). This standardized format ensures that AI applications can be easily integrated into clinical workflows while maintaining regulatory compliance. Its modular structure also enables scalability, making it ideal for institutions of all sizes.

Similarly, Kaapana[15], an open-source platform, provides a robust infrastructure for deploying AI workflows in clinical and collaborative settings. Its federated learning capabilities allow institutions to leverage AI models without transferring sensitive patient data, addressing privacy concerns while enabling broader collaboration. By integrating seamlessly with existing systems, Kaapana ensures that AI models can be deployed in a way that is both secure and effective.

The feasibility of such deployments is underscored by the increasing adoption of these frameworks in radiology and related fields. Their scalability and compatibility with existing healthcare infrastructures make them well-suited for theranostics, where the integration of AI could optimize decision-making and enhance patient outcomes. Additionally, these tools reduce the technical barriers to implementation, making it possible for institutions with varying levels of resources to adopt AI solutions.

To maximize the impact of the third ICPO call, it is essential to emphasize the importance of open-source frameworks for clinical deployment. By leveraging platforms like MONAI Deploy and Kaapana, researchers can not only test their models in real-world settings but also refine them based on clinical feedback. Moreover, this call should prioritize studies that demonstrate the clinical utility of AI algorithms, ensuring that the benefits of these technologies extend beyond research and into patient care.

By mandating clinical deployment and evaluation, the foundation ensures that supported projects address real-world challenges and deliver meaningful solutions. This approach fosters innovation and also paves the way for AI to become an integral part of clinical practice in nuclear medicine.

Conclusion

The feasibility of funding AI-driven studies in precision oncology and theranostics, as proposed by ICPO, is not only achievable but also timely and necessary. The challenges are clear—fragmented data, a lack of transparency in tool-sharing, and the gap between research and clinical implementation—but so are the solutions. This white paper underscores that with the right strategies and frameworks, these challenges can be addressed in a systematic and impactful way.

Data curation is the cornerstone of AI integration. Nuclear medicine has lagged behind radiology in data accessibility and standardization, but this gap can be closed. Advances in digital transformation, such as generative AI and open-source platforms like Gen3, have demonstrated their ability to streamline data aggregation and preparation, making this step feasible without being prohibitively resource-intensive. By prioritizing structured and collaborative approaches to data curation, ICPO's vision of creating a foundation for AI in theranostics is well within reach.

The application of AI algorithms to these curated datasets further cements the feasibility of ICPO's initiative. The success of open-source models like nnUNet, TotalSegmentator, and Enhance.PET illustrates that collaboration and transparency can fast-track progress, even in fields as niche as nuclear medicine. The open-sourcing of AI tools and models ensures that the benefits of these efforts will extend far beyond individual projects, fostering an ecosystem of shared innovation that is crucial for the field's growth.

Finally, the deployment of AI algorithms into clinical practice is not a distant aspiration—it is an achievable goal with existing frameworks like MONAI Deploy and Kaapana. These tools make it practical to integrate AI models into healthcare systems, enabling real-world validation of their utility. The emphasis on clinical deployment in ICPO's funding call ensures that supported projects go beyond theoretical promise and deliver measurable impact in improving patient outcomes - which is necessary.

ICPO's structured funding call, with its three interconnected pillars, represents a feasible and actionable strategy for advancing AI in precision oncology and theranostics. The infrastructure, tools, and collaborative practices needed to make this vision a reality are already available. By aligning its funding efforts with these resources, ICPO can bring forth meaningful progress in the field.

References

1. Clark, K. *et al.* The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J. Digit. Imaging* **26**, 1045–1057 (2013).
2. Marcus, D. S., Fotenos, A. F., Csernansky, J. G., Morris, J. C. & Buckner, R. L. Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. *J. Cogn. Neurosci.* **22**, 2677–2684 (2010).
3. Kandarpa, K. & Liu, G. Medical Imaging and Data Resource Center. https://www.nibib.nih.gov/programs/medical-imaging-data-resource-center.
4. Siragusa, I., Contino, S., La Ciura, M., Alicata, R. & Pirrone, R. MedPix 2.0: A Comprehensive Multimodal Biomedical Dataset for Advanced AI Applications. *arXiv [cs.DB]* (2024).
5. Markiewicz, C. J. *et al.* The OpenNeuro resource for sharing of neuroscience data. *Elife* **10**, (2021).
6. Saenz, A., Chen, E., Marklund, H. & Rajpurkar, P. The MAIDA initiative: establishing a framework for global medical-imaging data sharing. *Lancet Digit. Health* **6**, e6–e8 (2024).
7. EIC Pathfinder Challenges 2025. *European Innovation Council* https://eic.ec.europa.eu/eic-funding-opportunities/eic-pathfinder/eic-pathfinder-challenges-2025_en.
8. Gen3 Software —. *Center for Translational Data Science* https://ctds.uchicago.edu/gen3.
9. Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. *arXiv [cs.LG]* (2019).
10. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
11. Wasserthal, J. *et al.* TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiol Artif Intell* **5**, e230024 (2023).
12. Jorge Cardoso, M. *et al.* MONAI: An open-source framework for deep learning in healthcare. *arXiv [cs.LG]* (2022) doi:10.48550/ARXIV.2211.02701.
13. Sundar LS, Badawi RD, Spencer BA, Li E, Cherry SR, Abdelhafez YG, Hacker M, Jones T, Beyer T. Enhance-PET: exploring the human functional connectome using total-body [18F] FDG-PET.

European Journal of Nuclear Medicine and Molecular Imaging **48**, S201–S201 (2021).

14. Bericat, D. Taking AI into Clinical Production with MONAI Deploy. *NVIDIA Technical Blog* https://developer.nvidia.com/blog/taking-ai-into-clinical-production-with-monai-deploy/ (2022).
15. Kulla, L., Schader, P., Maier-Hein, K. & Nolden, M. Harmonized import of clinical research data for the open source image analysis platform kaapana. in *Bildverarbeitung für die Medizin 2024* 102–107 (Springer Fachmedien Wiesbaden, Wiesbaden, 2024).