Journal of Nuclear Medicine, published on December 15, 2022 as doi:10.2967/jnumed.121.263703

Artificial Intelligence in Nuclear Medicine: Opportunities, Challenges, and Responsibilities Toward a Trustworthy Ecosystem

Babak Saboury[1], Tyler Bradshaw[2], Ronald Boellaard[3], Irène Buvat[4], Joyita Dutta[5], Mathieu Hatt[6], Abhinav K. Jha[7], Quanzheng Li[8], Chi Liu[9], Helena McMeekin[10], Michael A. Morris[1], Peter J.H. Scott[11], Eliot Siegel[12], John J. Sunderland[13], Neeta Pandit-Taskar[14], Richard L. Wahl[15], Sven Zuehlsdorff[16], Arman Rahmim[17]

[1] Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, USA
[2] Department of Radiology, University of Wisconsin - Madison, USA
[3] Department of Radiology & Nuclear Medicine, Cancer Centre Amsterdam, Amsterdam University Medical Centres, Netherlands
[4] Institut Curie, Université PSL, Inserm, Université Paris-Saclay, Orsay, France
[5] Department of Electrical and Computer Engineering, University of Massachusetts Lowell, USA
[6] LaTIM, INSERM, UMR 1101, University of Brest, Brest, France
[7] Department of Biomedical Engineering and Mallinckrodt Institute of Radiology, Washington University in St. Louis, USA
[8] Department of Radiology, Massachusetts General Hospital and Harvard Medical School, USA
[9] Department of Radiology and Biomedical Imaging, Yale University, USA
[10] Department of Clinical Physics, Barts Health NHS Trust, United Kingdom
[11] Department of Radiology, University of Michigan Medical School, USA
[12] Department of Radiology and Nuclear Medicine, University of Maryland Medical Center, USA
[13] Departments of Radiology and Physics, University of Iowa, USA
[14] Department of Radiology, Memorial Sloan Kettering Cancer, USA
[15] Mallinckrodt Institute of Radiology, Washington University in St. Louis, USA
[16] Siemens Medical Solutions USA, Inc., Hoffman Estates, USA
[17] Departments of Radiology and Physics, University of British Columbia, Canada

Running title: Toward a Trustworthy AI Ecosystem

Word Count: 6459 (including abstract, text, table text, figure legends, references)

Financial support: None (individual COIs for each author are listed in the Disclosures section).

Keywords: Artificial intelligence, Trustworthy, Nuclear Medicine, Ecosystem

Corresponding Author:
Babak Saboury, MD, MPH, DABR, DABNM
Oncoradiologist and Nuclear Medicine Physician
Chief Clinical Data Science Officer
Lead Physician, Center for Precision Dosimetry and Radiopharmaceutical Treatment Planning
Department of Radiology and Imaging Sciences
Clinical Center, National Institutes of Health

Email: sabouryb@nih.gov
ORCID: 0000-0002-8632-6292
Phone: 240-507-3077
Fax: 301-480-5014

ABSTRACT

Trustworthiness is a core tenet of medicine. The patient-physician relationship is evolving from a dyad to a broader ecosystem of healthcare. With the emergence of artificial intelligence (AI) in medicine, the elements of trust must be revisited. We envision a roadmap for the establishment of trustworthy AI ecosystems in nuclear medicine. In this report, AI is contextualized in the history of technological revolutions. Opportunities for AI applications in nuclear medicine related to diagnosis, therapy and workflow efficiency, as well as emerging challenges and critical responsibilities are discussed. Establishing and maintaining leadership in AI requires a concerted effort to promote the rational and safe deployment of this innovative technology by engaging patients, nuclear medicine physicians, scientists, technologists, referring providers, among other stakeholders, while protecting our patients and society. This strategic plan is prepared by the AI Task Force of the Society of Nuclear Medicine and Molecular Imaging (SNMMI).

NOTEWORTHY

- An appropriate AI ecosystem can contribute to enhancing the trustworthiness of AI tools throughout their life-cycle through close collaboration among stakeholders (page 8).
- A "trustworthy medical AI system" depends upon the trustworthiness of the AI system itself, as well as the trustworthiness of all people and processes that are part of the system's life-cycle (page 15).
- By encouraging the establishment of trustworthy AI in nuclear medicine, SNMMI aims to decrease health disparity, increase health system efficiency, and contribute to the improved overall health of society using AI applications in the practice of nuclear medicine (page 22).

INTRODUCTION

Medicine utilizes science, practical wisdom, and the best available tools in the art of compassionate care. The necessity of dealing with maladies has motivated physicians to incorporate inventions into medical practice to decrease or eliminate patient suffering. During the past two centuries, along with technological revolutions, new medical devices have become the standard of care, from the stethoscope and electrocardiogram to cross-sectional imaging (Figure 1). The stethoscope, which arose out of the first industrial revolution, is so pervasive that it has become the symbol of healthcare professionals today. Compared to other medical equipment, it has the highest positive impact on the perceived trustworthiness of the practitioner seen with it (*1*).

Nuclear medicine has always embraced the progress of technology. With the emergence of AI we will again be poised to experience a modern renaissance, similar to the one experienced following David Kuhl's and Roy Edwards' groundbreaking work in the 1960s. By applying the concepts of Radon transform through newly available computing technology, they introduced volumetric cross-sectional medical imaging with single-photon emission computed tomography (SPECT), which was subsequently followed by the development of X-ray based computed tomography (CT) and positron emission tomography (PET) (*2*).

The past decades have seen tremendous advances in information technology and in its integration into the practice of medicine. The application of artificial intelligence (AI) to medicine represents the actualization of a new era. Such transformative technologies can affect all facets of society, yielding advances in space exploration, defense, energy, industrial processes, finance, even cartography, transportation, and food service among others.

The addition of AI into clinical practice in nuclear medicine poses opportunities and challenges. The full benefits of this new technology will continuously evolve. It is important to recognize that the nuclear medicine community must be actively involved to ensure safe and effective implementation. Establishing and maintaining AI leadership in the realm of nuclear medicine requires a comprehensive strategy to promote the application of innovative technology, while protecting our patients and society, executing our professional and ethical obligations, and promoting our values. A potential advantage of deploying AI techniques is that nuclear medicine methodologies may become more widely available, increasing access of patients to high quality nuclear medicine procedures.

Nuclear medicine professional societies such as the Society of Nuclear Medicine and Molecular Imaging (SNMMI) and others provide leadership to ensure we recognize the benefits of technological advances in a manner consistent with our core values, medical ethics and society's best interests. In July 2020, the SNMMI formed the AI Task Force by bringing together experts in nuclear medicine and AI, including physicists, computational imaging scientists, physicians, statisticians, and representatives from industry & regulatory agencies. This article serves as both a strategic plan and summary of the deliberations of the SNMMI AI Task Force over the past year in conjunction with other focused topics, including best practices for development (*3*) and evaluation (*4*) (Table 1).

OPPORTUNITIES

Quantitative Imaging and Process Improvement

Nuclear medicine is evolving toward even better image quality and more accurate and precise quantification in the precision medicine era, most recently in the paradigm of theranostics.

Diagnostic Imaging

AI techniques in the “patient-to-image subdomain” improve acquisition and models in the “image-to-patient subdomain” enable improved decision making for interventions on patients (Figure 2) (*3*).

Image generation considerations are elaborated in supplement opportunities part A, however examples include improved image reconstruction from raw data (list-mode, sinogram), data corrections including for attenuation, scatter and motion, and post-reconstruction image enhancement, among others (*5*–*7*). These enhancements could impact PET and SPECT in clinical use today. Multi-time-point acquisitions and PET/MR may see improved feasibility.

Specific opportunities in *image analysis* are elaborated in the supplement (opportunities part B). A few examples include image registration, organ and lesion segmentation, biomarker measurements and multi-omics integration, and kinetic modeling (*8*).

Opportunities for clinical utilization of AI in nuclear medicine practice were extensively reviewed recently, including brain imaging (*9*), head and neck imaging (*10*), lung imaging (*11*), cardiac imaging (*12*,*13*), vascular imaging (*13*,*14*), bone imaging (*15*), prostate imaging (*16*), and imaging of lymphoma (*17*). Neuroendocrine tumors, other cancers (including, gastrointestinal, pancreas, hepatobiliary, sarcoma, and hereditary cancers), infection and inflammation are some examples of additional areas requiring further consideration.

Emerging Nuclear Imaging Approaches

New developments are also emerging such as Total Body PET (TB-PET) (*18*) which presents unique data and computational challenges. Another potential use of AI is to separate multi-channel data from single session multi-isotope dynamic PET imaging. This pragmatic advancement could be valuable to extract greater phenotyping information in evaluation of tumor heterogeneity (*19*).

Radiopharmaceutical Therapies (RPTs)

There are several areas where AI is expected to significantly impact RPTs:

AI-Driven Theranostic Drug Discovery and Labeling. The use of AI for molecular discovery has been explored to select the most promising leads to design suitable theranostics for the target in question. For example, machine learning models could be trained using parameters from past theranostic successes and failures (e.g. logP, K_d, BP) to establish which best predict a given outcome (e.g. specific binding, blood-brain-barrier penetration, tumor to muscle ratio). New AI approaches are revolutionizing our understanding of protein-ligand interactions (*20*). New hit molecules (e.g. from the literature or high throughput screens) can then serve as the test set in such AI models to speed up hit-to-lead optimization. Subsequently, with lead molecules identified, AI could also predict optimal labeling precursors and synthesis routes to facilitate fast and efficient development of theranostic agents (*21*,*22*). By defining parameters from existing synthetic datasets (e.g. solvents, additives, functional groups, NMR shifts) models can be trained to predict radiochemical yield for a given substrate using different precursors and radiosynthetic methods. Subjecting new lead candidates as test sets in the models will enable rapid identification of appropriate precursors and labeling strategies for new theranostics minimizing resource intensive manual synthetic development.

Precision Dosimetry. The field of radiopharmaceutical dosimetry is progressing rapidly. After administration of radiopharmaceuticals, dynamical and complex pharmacokinetics results in time-variable biodistribution. Interaction of ionizing particles arising from the injected agent with the target and normal tissue results in energy deposition. Quantification of this deposited energy and its biological effect is the essence of dosimetry with opportunities to link the deposited energy to its biological effect on diseased and normal tissues (Figure 3).

In dosimetry, SPECT serves as a post-treatment quantitative measuring device. One challenge is the difficulty for patients to remain flat and motionless on the scanning table for the required time. AI-based image reconstruction/enhancement methods can: reduce required SPECT scan time for patients while maintaining or enhancing accuracy of quantification (*23*), and enable attenuation correction in SPECT (*24*).

Multiple steps in dosimetry potentially can be enhanced by AI methods, including multi-modality and multi-time point image registration, segmentation of organs and tumors, time activity curve (TAC) fitting, time-integrated activity (TIA) estimation, conversion of TIA into absorbed dose, linking macro-scale dosimetry to micro-scale dosimetry and arriving at comprehensive patient dose profiling (*25*).

Predictive Dosimetry and Digital Twins. Existing models can perform dosimetry before (e.g., I-131 MIBG) or following treatment. Personalized RPTs require predictive dosimetry for optimal dose prescription where AI can play a role. Pre-therapy (static or dynamic) PET scans could model radiopharmaceutical pharmacokinetics and absorbed doses in tumors and normal organs. Furthermore, it is possible to additionally utilize intra-therapy scans (e.g. single-time-point SPECT in the first cycle of RPTs) to better anticipate and adjust doses in subsequent cycles.

Overall, a vision of the future involves accurate and rapid evaluation of different RPT approaches (e.g. varying injected radioactivity dose and rate, site of injection, injection intervals, coupling with other therapies, etc.) utilizing the concept of the theranostic digital twin (TDT). The theranostic digital twin can aid nuclear medicine physicians in complex decision-making processes. It enables experimentation (in the digital world) with different treatment scenarios, thus optimizing delivered therapies.

The opportunities discussed in the radiopharmaceutical therapy section above are further described in supplement opportunities part C.

Clinical Workflow: Increase Throughput While Maintaining Excellence

AI may impact operations in nuclear medicine, such as patient scheduling and resource utilization (*26*), predictive maintenance of devices to minimize unexpected downtimes, monitoring quality control measurement results to discover hidden patterns and indicate the potential for improvement, and monitoring performance of devices in real-time to capture errors and detect aberrancies (*26*,*27*). These processes will make the practice of nuclear medicine safer, more reliable, and more valuable.

Triage of urgent findings and augmentation of time-consuming tasks could improve report turn-around-time of the most critical cases and increase the efficiency of nuclear medicine physicians allowing them to more effectively care for patients. It is important to assure AI systems in nuclear medicine are sustainable through new current procedural terminology (CPT) code development and appropriate relative value unit (RVU) assignment for the technical and professional components. It is also possible that increased efficiencies in interpretation (more cases read accurately per unit time) may allow AI to be deployed into clinical workflows in an overall cost-effective manner.

AI ECOSYSTEM

Actualization of Opportunities and Contextualization of Challenges

While early nuclear medicine AI systems are already emerging, many opportunities remain in which the continuous propagation of AI technology could augment our precision patient care and practice efficiencies. The environment where AI development, evaluation, implementation, and dissemination occurs needs a sustainable ecosystem to enable progress, while appropriately mitigating concerns of stakeholders.

The total life-cycle of AI systems, from concept to appropriation of training-data, model development and prototyping, production testing, validation and evaluation, implementation/deployment, and post-deployment surveillance, occurs within a framework which we call the *"AI Ecosystem"* (Figure 4). An appropriate AI ecosystem can contribute to enhancing the trustworthiness of AI tools throughout their life-cycle through close collaboration among stakeholders.

CHALLENGES FOR DEVELOPMENT, VALIDATION, DEPLOYMENT, AND IMPLEMENTATION

Development of AI Applications/Medical Devices

Five challenges that should be addressed include data, optimal network architecture, measurement and communication of uncertainty, identification of clinically impactful use cases, and improvements in team science approaches (supp Development Challenges).

Evaluation (Verification of Performance)

Theories on appropriate evaluation of AI software are a broad and active area of current investigation. Establishing clear and consistent guidelines for performance profiling remains challenging. Most current verification studies evaluate AI methods based on metrics that are agnostic to performance in clinical tasks (28). While such evaluation may help demonstrate promise, there is an important need for further testing on specific clinical tasks before the algorithms can be implemented. Failure mode profiling is among the most important challenges (supp Evaluation Challenges).

Ethical, Regulatory, and Legal Ambiguities

Major ethical concerns include informed consent for data usage, replication of historical bias and unfairness embedded in training data, unintended consequences of AI device agency, inherent opaqueness of some algorithms, concerns about the impact of AI on healthcare disparities, and trustworthiness (supp Ethical Challenges). AI in nuclear medicine has limited legal precedent (29).

Implementation of Clinical AI Solutions & Post-Deployment Monitoring

The lack of an AI-Platform integrating AI applications in nuclear medicine workflow in among the most critical challenges of implementation (30). Barriers of dissemination can be categorized at individual level (healthcare providers), at the institutional level (organization culture), and at the societal level (31). Deployment is not the end of the implementation process. (supp Implementation Challenges).

TRUST AND TRUSTWORTHINESS

In medicine, trust is the essence, not a pleasance.

Successful solutions to the above-mentioned challenges are necessary but not sufficient for the sustainability of AI ecosystems in medicine. Well-developed and validated AI devices with supportive regulatory context, appropriate reimbursement and successful primary implementation may still fail if physicians, patients, and society lose trust due to lack of transparency and other critical elements of trustworthiness such as perceived inattention to health disparity or racial injustice. In a recent survey, Martinho et al. (*32*) found significant perceived mistrust among healthcare providers in regard to AI systems and the AI industry while realizing the importance and benefits of this new technology. Responders also emphasized the importance of ethical utilization, and the need for physician-in-the-loop interactions with AI systems, among the other factors. There is a need for a comprehensive analysis of the AI ecosystem to define and clarify the core elements of trustworthiness in order to realize the benefits of AI in clinical practice.

RESPONSIBILITIES - TOWARD TRUSTWORTHY AI

When the safety, wellbeing, and rights of our patients are at stake, SNMMI should be committed to support principles that are future-proof and innovation-friendly.

The willingness of physicians and patients to depend on a specific tool in a risky situation is the measure of “trustworthiness” of that tool (*33*). In the case of AI systems, that willingness is based on a set of specific beliefs about reliability, predictability, and robustness of the tool as well as integrity, competency and benevolence of the people/processes involved in the AI system’s life-cycle (development, evaluation/validation, deployment/implementation, and use).

A “trustworthy medical AI system” depends upon the trustworthiness of the AI system itself, as well as the trustworthiness of all people and processes that are part of the system’s life-cycle (Figure 5).

Trustworthy medical AI systems require a societal and professional commitment to the ethical AI framework which includes four principles rooted in the fundamentals of medical ethics. These principles should be observed in various phases of the AI system life-cycle: respect for patients’ and physicians’ autonomy, prevention of harm, beneficence to maximize the wellbeing of patients and society, fairness.

In what follows, we outline twelve key elements that need to be consistently present in AI systems.

12 Key Elements of Trustworthy AI Systems

Human Agency. AI systems should empower physicians and patients, allowing them to make better informed decisions and foster their autonomy *(34)*. Effects of the AI algorithms on human independence should be considered. It should be clear to patients and physicians the extent to which AI is involved in patient care and the extent of physician oversight. There must be checks to avoid automation bias, which is the propensity of humans to value and overly rely on observations and analyses from computers over those of human beings (*35*).

Oversight. There must be sufficient oversight of AI decision-making, which can be achieved through human-in-the-loop, and human-in-command approaches (36). AI systems that are involved in higher-risk tasks (e.g. drives clinical management, diagnose, or treat disease) must be closely monitored through post-market surveillance by independent professional credentialing organizations analogous to certification and recertification of medical professionals. Peer review processes in practices can be adapted to consider the combined physician/AI decision making process.

Technical Robustness. AI systems must perform in a dependable manner (sufficient accuracy, reliability, and reproducibility) (*37*). This performance should be resilient to the breadth of clinical circumstances related to their prescribed use (generalizability). The AI tool should explicitly convey a degree of certainty about its output (confidence score) and have a mechanism in place to monitor the accuracy of outputs as part of a continuous quality assurance program. Failure modes of the algorithm should be well characterized, documented, and understood by users.

Safety & Accountability. According to the concepts of safety-critical systems *(38)*, AI systems should prioritize safety above other design considerations (e.g. potential gains in efficiency, economics, or performance). When adverse events occur, mechanisms should be in place for ensuring accountability and redress. Vendors must be accountable for the claims made of their AI systems. Physicians must be accountable for the way in which AI systems are implemented and used in the care of their patients. The ability to independently audit the root cause of a failure in an AI system is important. Protection must be provided for individuals or groups reporting legitimate concerns in accordance with the principles of risk management.

Security and Data Governance. AI systems must include mechanisms to minimize harm as well as to prevent it wherever possible. They must comply with all required cybersecurity standards. There should be an assessment of vulnerabilities such as data poisoning, model evasion, and model inversion. Assurances should be made to mitigate potential vulnerabilities and avoid misuse, inappropriate use, or malicious use (such as a deep fake) (39).

Predetermined Change Control Plan. AI tools can be highly iterative and adaptive in nature which may lead to rapid continual product improvement. The plan should include types of anticipated modifications (Software as a Medical Device Pre-Specifications). There must be a clear and well-documented methodology (algorithm change protocol) to evaluate the robustness and safety of the updated AI system. The algorithm change protocol should include guidelines for data management, re-training, performance evaluation, and update procedures. Vendors should maintain a culture of quality and organizational excellence.

Diversity, Bias-awareness, Non-discrimination, and Fairness. AI systems can be affected by input data maladies (incomplete data, inadvertent historically biased data), algorithm design insufficiencies, or suboptimal performance assessment/monitoring strategies. These issues may result in biases leading to unintended prejudice and cause harm to patients. Discriminatory bias should be removed from AI systems in the development phase where possible (31).

AI system performance should be evaluated in a wide spectrum of diseases and in patients suffering from a particular condition regardless of extraneous personal characteristics. No particular group of patients should be systematically excluded from AI device development. Patients who are underrepresented or suffer from rare diseases should not be excluded from AI systems development or evaluation—though such datasets will be sparse and most likely could only be used in the evaluation of AI methods developed in larger populations (for generalizability). Appropriate validation testing on standardized sets that incorporate patient diversity, including rare or unusual presentation of disease, are critical to evaluate the presence of bias in results regardless of the training data used (40).

AI systems should be user-centric and developed with an awareness of the practical limitations of the physician work environment in mind. Accessibility features should be provided to those individuals with disabilities to the extent necessary according to *universal design* *principles*.

Stakeholder Participation. Throughout the life-cycle of an AI system, all the stakeholders that may directly or indirectly be affected should actively participate to help, advise and oversee the developers and industry. Participation of patients, physicians and all the relevant providers, healthcare systems, payors, regulatory agencies, and professional societies is imperative. This inclusive and transparent engagement is essential for a Trustworthy AI ecosystem. Regular clinical feedback is needed to establish longer term mechanisms for active engagement.

Transparency & Explainability. Vendors should provide open communication of how an AI system is validated for the labeled claim (purpose, criteria, and limitations) by describing the clinical task for which the algorithm was evaluated, composition of the patient population used for validation, the image acquisition, reconstruction and analysis protocols, and the figure of merit used for the evaluation (*4*,*37*). There must be appropriate training material and disclaimers for healthcare professionals on how to adequately use the system. It should be clear which information is communicated from the AI system and which information is communicated by a healthcare professional. AI systems should incorporate mechanisms to log and review which data, AI model, or rules were used to generate certain outputs (auditability and traceability). The effect of the input data on the AI system's output should be conveyed in a manner whereby their relationship can be understood by physicians and *ideally* patients (explainability) in order to allow a mechanism to critically evaluate and contest the AI system outputs. For diagnostic applications, the AI system should communicate a degree of confidence (uncertainty) together with its decision. To the extent possible, in high stakes tasks the use of 'black box' AI systems without proper emphasis on transparency should be avoided (41).

Sustainability of Societal Wellbeing. It is important to acknowledge that exposure to AI could negatively impact social relationships and attachment within the healthcare system (social agency) (42). AI systems should be implemented in a manner that enhances the physician-patient relationship. AI systems should not interfere with human deliberation or deteriorate social interactions. The societal and environmental impact of an AI tool should be carefully considered to ensure sustainability. Healthcare workers who are impacted by the implementation of AI systems should be given an opportunity to provide feedback and contribute to its implementation plan. Steps should be taken by professional societies and training programs to assure AI systems do not result in de-skilling of professionals, such as providing opportunities for re- and up-skilling. Rather, a new set of skills including physician oversight and interaction with AI tools, will evolve and must be refined.

Privacy. AI systems should have appropriate processes in place to maintain security and privacy of patient data. A minimal amount of personal data required should be utilized (data minimization). There should be a statement on measures employed to achieve *privacy-by-design* such as encryption, pseudoanonymization, aggregation, and anonymization. Systems should be aligned with standards and protocols for data management and governance.

Fairness and Supportive Context of Implementation. Early development efforts can pose more risk to developers and consumers. In order to address the concerns related to liability, there have been successful programs in other industries to encourage adoption of new technology and support consumer protection such as for vaccines and autonomous vehicles (29).

STRATEGIES FOR SUCCESS

Part 1: SNMMI Initiatives

In July 2022 SNMMI created the AI Task Force to strategically assess the emergence of AI in nuclear medicine (supp SNMMI initiatives). Areas of important focus were designated working groups, such as the AI & Dosimetry working group for predictive dosimetry and treatment planning.

PART 2: SNMMI Action Plan

The Task Force recommends the establishment of the AI Center of Excellence (AICE) to facilitate a sustainable AI ecosystem (supp SNMMI action plan). A Nuclear Medicine Imaging Archive (NMIA) will address the need for meaningful data access. The Trustworthy AI in Medicine and Society Coalition (TAIMS coalition) will address the need for an AI Bill of Rights *(43)*.

Part 3: SNMMI Recommendations

Recommendations for the future are also provided in the supplement (supp SNMMI Recommendations).

CONCLUSION

There are immense and exciting opportunities for AI to benefit the practice of nuclear medicine. Meanwhile, there are challenges that must and can be addressed head-on. As current challenges are addressed and new AI solutions emerge, SNMMI and the nuclear medicine community have the responsibility to ensure trustworthiness of these tools in the care of patients.

We can all benefit from efforts to ensure fairness, inclusion, and lack of bias in the entire lifecycle of AI algorithms in different settings.

There are three levels of facilitation that can support and enable the appropriate environment for trustworthy AI. First, our community must establish guidelines, such as those referenced in this article, to promote the natural development of trustworthy AI. Second, we can faciliate trustworthy AI through an AI Center of Excellence (AICE). Third, we can make trustworthy AI occur through active engagement and communicative actions.

By encouraging establishment of trustworthy AI in nuclear medicine, SNMMI aims to decrease health disparity, increase health system efficiency, and contribute to the improved overall health of society using AI applications in the practice of nuclear medicine.

ACKNOWLEDGMENTS

The members of the Task Force wish to acknowledge Bonnie Clarke for her very valuable support of this project. We are grateful for the insightful comments of Dr. Roberto Maas-Moreno, particularly on the concept of theranostic digital twins (AI and radiopharmaceutical therapy).

DISCLOSURES

The views expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. government nor do they reflect any official recommendation or endorsement of the National Institutes of Health (NIH). Helena McMeekin is a part-time employee of Hermes Medical Solutions, Inc. Sven Zuehlsdorff is a full-time employee of Siemens Medical Solutions, Inc. No other potential conflicts of interest relevant to this article exist.

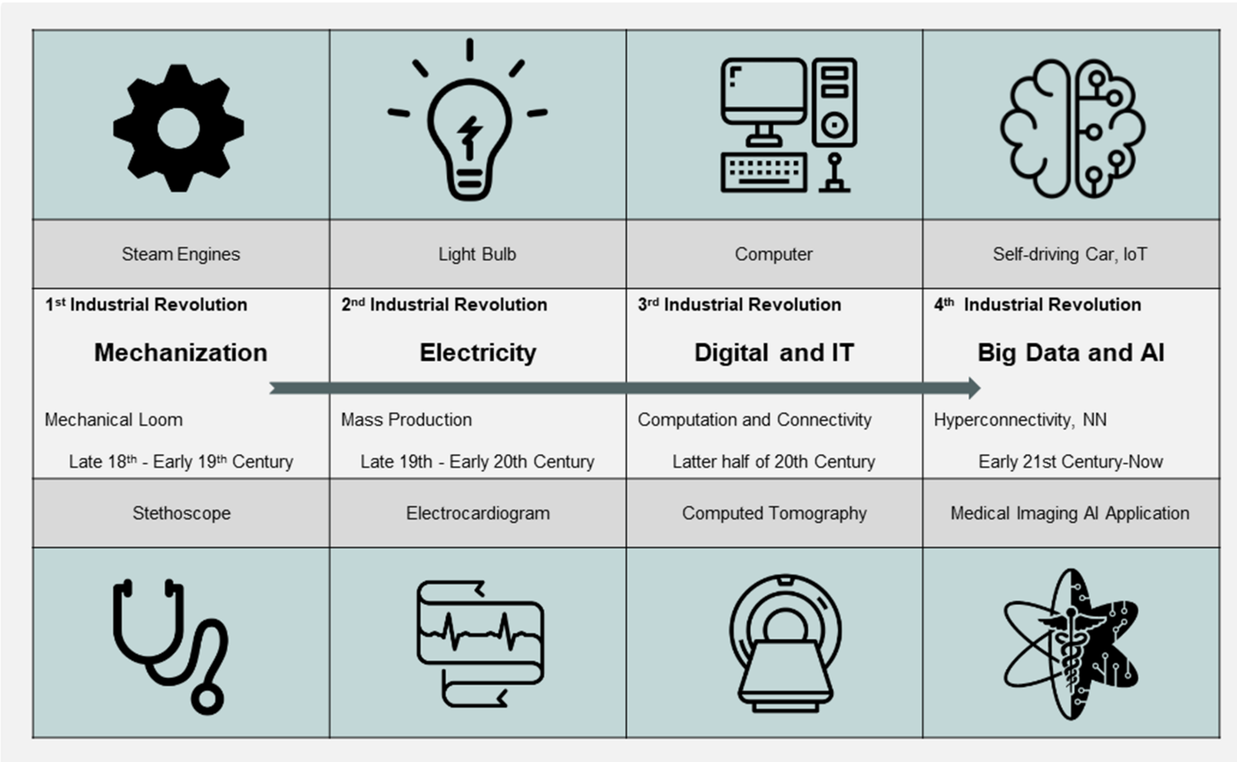

Figure 1

New technologies in Medicine have coincided with each phase of the industrial revolution. The first industrial revolution was mechanization with the mechanical loom invented in 1784. The stethoscope was invented by René Laennec in 1816 and improved upon by Arthur Leared (1851) and George Philip Cammann (1852). The second industrial revolution was driven by the advent of electricity with the commercial light bulb patented by Thomas Edison in 1879, the telegram, and the modern factory production line. The electrocardiogram was invented by Augustus Waller in 1887 by projecting the heartbeat captured by a lippmann capillary electrometer onto a photographic plate allowing the heartbeat to be recorded in real-time. Willem Einthoven (1895) assigned the letters P, Q, R, S, and T to the theoretical waveform. The third industrial revolution, known as the digital revolution, brought computing technology and refined it to the personal computer. In the 1960's, Kuhl and Edwards developed cross-sectional computed tomography and implemented this in the SPECT scanner, which was later applied to the CT scanner by Sir Godfrey Hounsfield and Allan Cormack in 1972. The fourth industrial revolution is the modern day with Big Data, hyperconnectivity, and neural networks (NN) resulting in the ability to propel self-driving cars and the development of AI in nuclear medicine.

Table 1

Opportunities and Challenges Ahead of Nuclear Medicine Toward Achieving Trustworthy AI: an Outline for Discussion

Category	Domain	Sub-domain
Opportunities		
	1. Diagnostic Imaging	
		A. Emerging Nuclear Imaging Approaches
	2. Radiopharmaceutical Therapies (RPTs)	
		A. AI-Driven Theranostic Drug Discovery and Labeling
		B. Precision Dosimetry
		C. Predictive Dosimetry and Digital Twins
	3. Clinical Workflow: Increase Throughput While Maintaining Excellence	
Challenges		
	1. Development of AI Applications/Medical Devices	
		A. Data
		B. Optimal Network Architecture
		C. Measurement and Communication of Uncertainty
		D. Clinically Impactful Use Cases
		E. Team Science
	2. Evaluation (Verification of Performance)	
		A. Performance Profiling Through Task-Based Evaluations
		B. Guidelines for Validation
		C. Multi-Center Clinical Trial Network
	3. Ethical, Regulatory, and Legal Ambiguities	
		A. Ethical Aspects
		B. Regulatory and Legal Aspects
	4. Implementation of Clinical AI Solutions & Post-Implementation Monitoring	
		A. AI-Platform
		B. Barriers of Dissemination and Implementation of AI Technology in Medicine
		C. Post-Deployment: Change Management & Performance
	5. Trust and Trustworthiness	

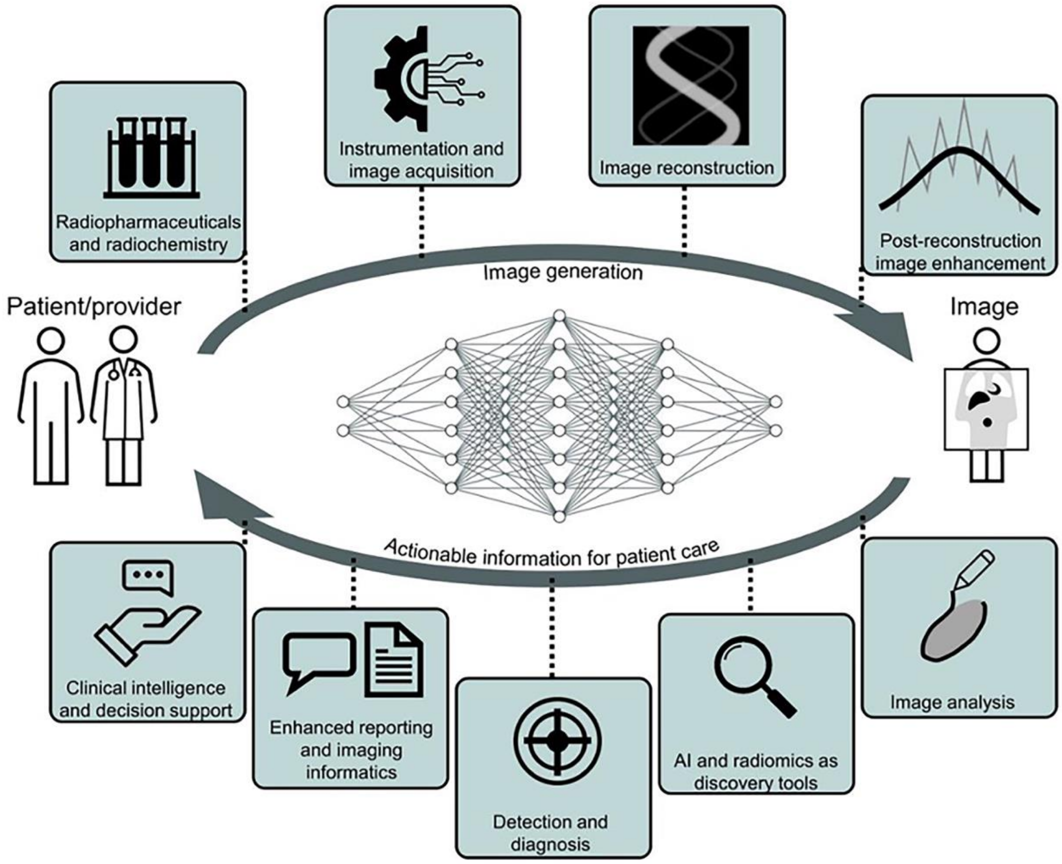

Figure 2

From patient to image creation and back to the physician, there are opportunities for AI systems to act at nearly any step in the medical imaging pipeline to improve our ability to care for patients and understand disease. (*3*)

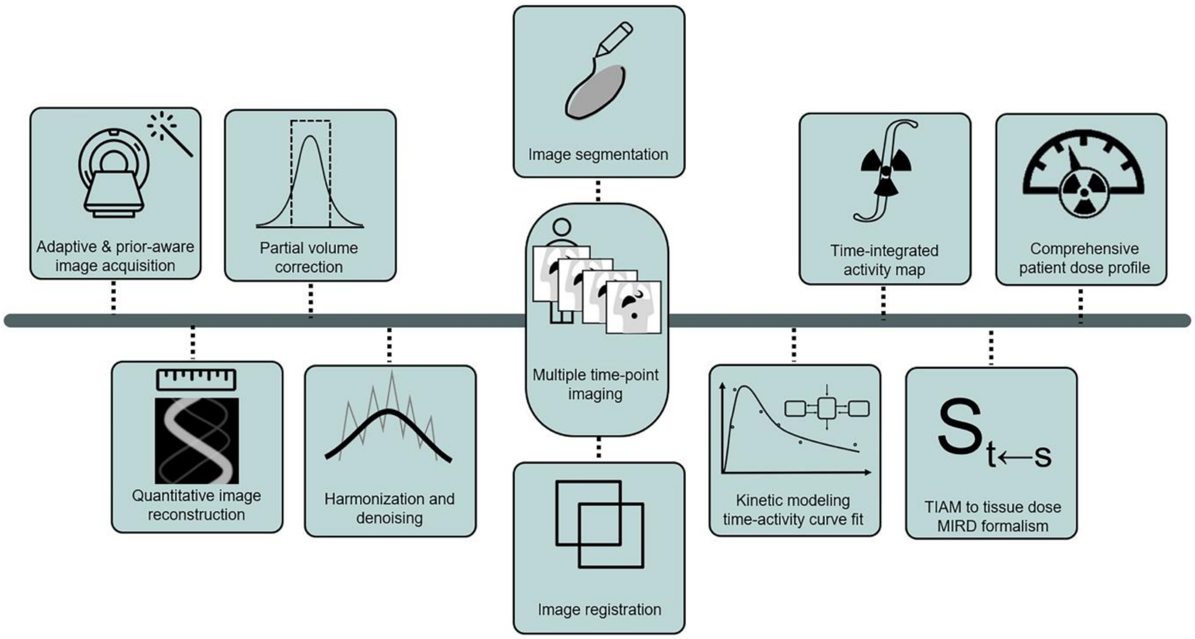

Figure 3

Dosimetry as a major frontier supported by AI towards personalization of therapy: Various contributions by AI to image acquisition, image generation and processing, followed by automated dose calculations can enable routine deployment and clinical decision support.

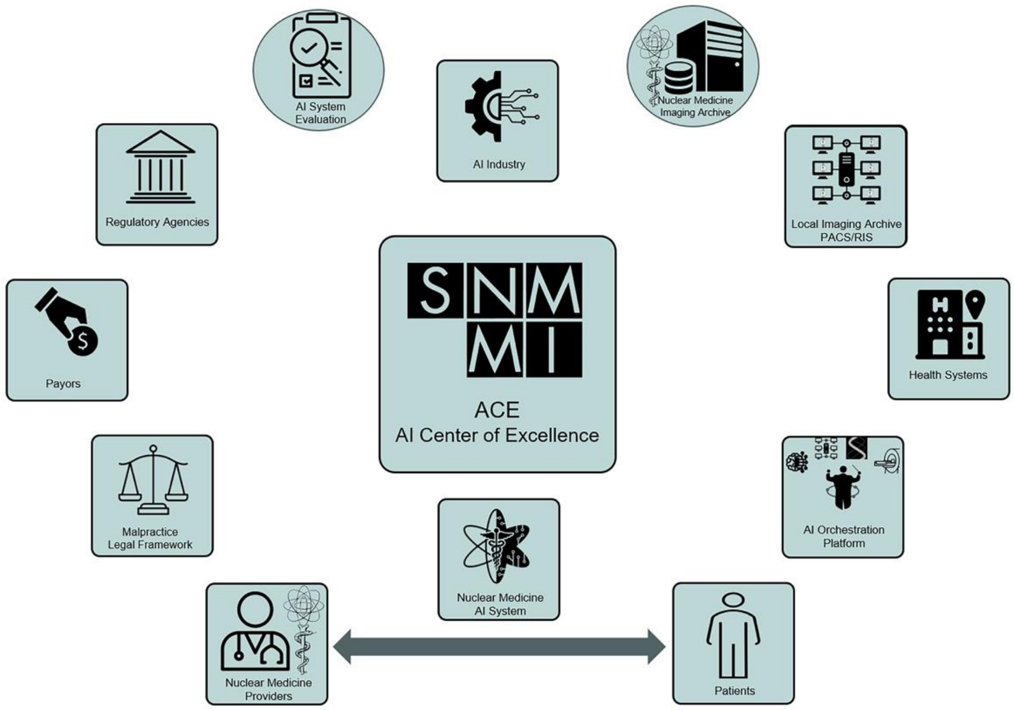

Figure 4

The AI Ecosystem is a complex system that represents the environment in which AI system development occurs. The Ecosystem connects stakeholders from industry to regulatory agencies, to physicians, to patients, to health systems and payers. The proposed SNMMI AI Center of Excellence (AICE) can serve as an honest broker to empower the AI Ecosystem from a neutral standpoint with a focus on solutions.

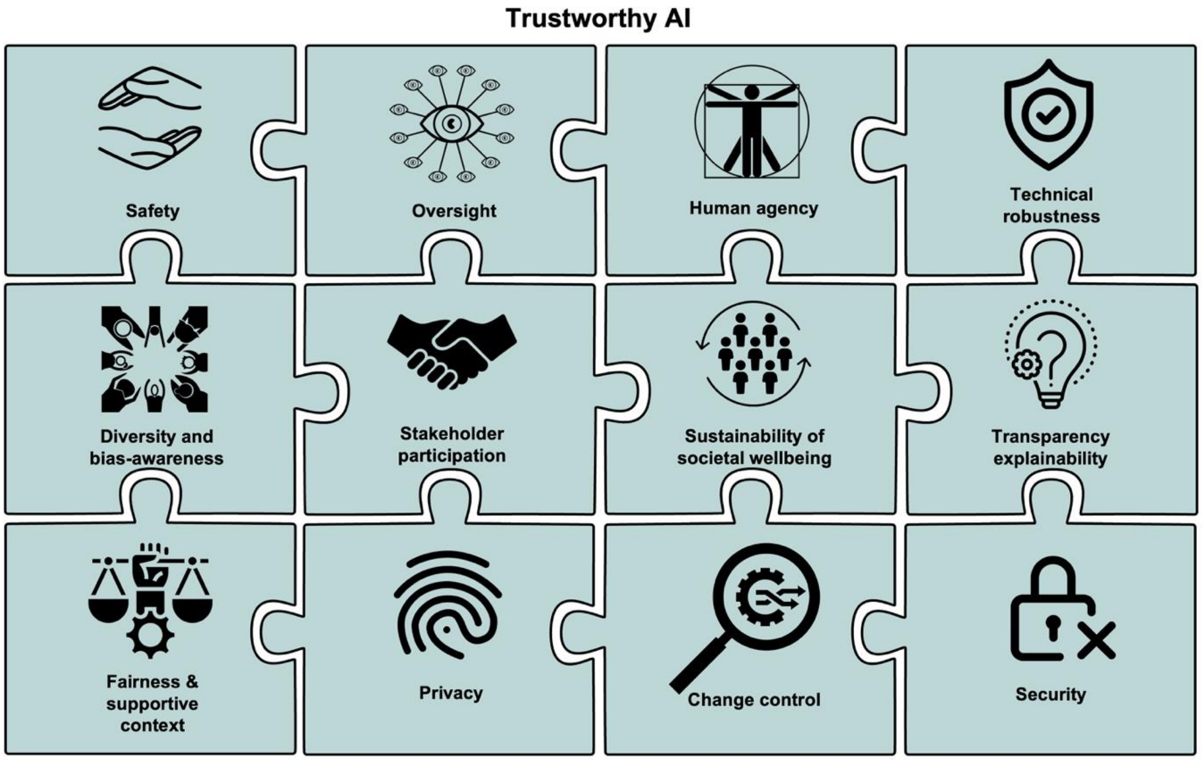

Figure 5

Twelve core concepts critical to Trustworthy AI Ecosystems

REFERENCES

1. Jiwa M, Millett S, Meng X, Hewitt VM. Impact of the presence of medical equipment in images on viewers' perceptions of the trustworthiness of an individual on-screen. *J Med Internet Res*. 2012;14:e100.

2. Dunnick NR. David E. Kuhl, MD. *Radiology*. 2017;285:1065.

3. Bradshaw TJ, Boellaard R, Dutta J, et al. Nuclear Medicine and Artificial Intelligence: Best Practices for Algorithm Development. *J Nucl Med*. 2022;63:500-510.

4. Jha AK, Bradshaw TJ, Buvat I, et al. Nuclear Medicine and Artificial Intelligence: Best Practices for Evaluation (the RELAINCE guidelines). *J Nucl Med*. 2022;63:1288-1299.

5. Gong K, Kim K, Cui J, Wu D, Li Q. The Evolution of Image Reconstruction in PET: From Filtered Back-Projection to Artificial Intelligence. *PET Clin*. 2021;16:533-542.

6. McMillan AB, Bradshaw TJ. Artificial Intelligence–Based Data Corrections for Attenuation and Scatter in Position Emission Tomography and Single-Photon Emission Computed Tomography. *PET Clin*. 2021;16:543-552.

7. Liu J, Malekzadeh M, Mirian N, Song T-A, Liu C, Dutta J. Artificial Intelligence-Based Image Enhancement in PET Imaging: Noise Reduction and Resolution Enhancement. *PET Clin*. 2021;16:553-576.

8. Yousefirizi F, Jha AK, Brosch-Lenz J, Saboury B, Rahmim A. Toward High-Throughput Artificial Intelligence-Based Segmentation in Oncological PET Imaging. *PET Clin*. 2021;16:577-596.

9. Cross DJ, Komori S, Minoshima S. Artificial Intelligence for Brain Molecular Imaging. *PET Clin*. 2022;17:57-64.

10. Gharavi SMH, Faghihimehr A. Clinical Application of Artificial Intelligence in PET Imaging of Head and Neck Cancer. *PET Clin*. 2022;17:65-76.

11. Zukotynski KA, Gaudet VC, Uribe CF, Chiam K, Bénard F, Gerbaudo VH. Clinical Applications of Artificial Intelligence in Positron Emission Tomography of Lung Cancer. *PET Clin*. 2022;17:77-84.

12. Miller RJH, Singh A, Dey D, Slomka P. Artificial Intelligence and Cardiac PET/Computed Tomography Imaging. *PET Clin*. 2022;17:85-94.

13. Slart RHJA, Williams MC, Juarez-Orozco LE, et al. Position paper of the EACVI and EANM on artificial intelligence applications in multimodality cardiovascular imaging using SPECT/CT, PET/CT, and cardiac CT. *Eur J Nucl Med Mol Imaging*. 2021;48:1399-1413.

14. Paravastu SS, Theng EH, Morris MA, et al. Artificial Intelligence in Vascular-PET: Translational and Clinical Applications. *PET Clin*. 2022;17:95-113.

15. Paravastu SS, Hasani N, Farhadi F, et al. Applications of Artificial Intelligence in 18F-Sodium Fluoride Positron Emission Tomography/Computed Tomography:: Current State and Future Directions. *PET Clin*. 2022;17:115-135.

16. Ma K, Harmon SA, Klyuzhin IS, Rahmim A, Turkbey B. Clinical Application of Artificial Intelligence in Positron Emission Tomography: Imaging of Prostate Cancer. *PET Clin*. 2022;17:137-143.

17. Hasani N, Paravastu SS, Farhadi F, et al. Artificial Intelligence in Lymphoma PET Imaging:: A Scoping Review (Current Trends and Future Directions). *PET Clin*. 2022;17:145-174.

18. Wang Y, Li E, Cherry SR, Wang G. Total-Body PET Kinetic Modeling and Potential Opportunities Using Deep Learning. *PET Clin*. 2021;16:613-625.

19. Ding W, Yu J, Zheng C, et al. Machine learning-based noninvasive quantification of single-imaging session dual-tracer 18F-FDG and 68Ga-DOTATATE dynamic PET-CT in oncology. *IEEE Trans Med Imaging*. 2021;PP.

20. Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. *Nature*. 2021;596:590-596.

21. Webb EW, Scott PJH. Potential Applications of Artificial Intelligence and Machine Learning in Radiochemistry and Radiochemical Engineering. *PET Clin*. 2021;16:525-532.

22. Ataeinia B, Heidari P. Artificial Intelligence and the Future of Diagnostic and Therapeutic Radiopharmaceutical Development:: In Silico Smart Molecular Design. *PET Clin*. 2021;16:513-523.

23. Arabi H, AkhavanAllaf A, Sanaat A, Shiri I, Zaidi H. The promise of artificial intelligence and deep learning in PET and SPECT imaging. *Phys Med*. 2021;83:122-137.

24. Shi L, Onofrey JA, Liu H, Liu Y-H, Liu C. Deep learning-based attenuation map generation for myocardial perfusion SPECT. *Eur J Nucl Med Mol Imaging*. 2020;47:2383-2395.

25. Brosch-Lenz J, Yousefirizi F, Zukotynski K, et al. Role of Artificial Intelligence in Theranostics:: Toward Routine Personalized Radiopharmaceutical Therapies. *PET Clin*. 2021;16:627-641.

26. Beegle C, Hasani N, Maass-Moreno R, Saboury B, Siegel E. Artificial Intelligence and Positron Emission Tomography Imaging Workflow: *PET Clin*. 17:29-37.

27. Ullah MN, Levin CS. Application of Artificial Intelligence in PET Instrumentation. *PET Clin*. 2022;17:175-182.

28. Yang J, Sohn JH, Behr SC, Gullberg GT, Seo Y. CT-less Direct Correction of Attenuation and Scatter in the Image Space Using Deep Learning for Whole-Body FDG PET: Potential Benefits and Pitfalls. *Radiol Artif Intell*. 2021;3:e200137.

29. Mezrich JL. Demystifying Medico-legal Challenges of Artificial Intelligence Applications in Molecular Imaging and Therapy. *PET Clin*. 2022;17:41-49.

30. Saboury B, Morris M, Siegel E. Future Directions in Artificial Intelligence. *Radiol Clin North Am*. 2021;59:1085-1095.

31. Yousefi Nooraie R, Lyons PG, Baumann AA, Saboury B. Equitable Implementation of Artificial Intelligence in Medical Imaging: What Can be Learned from Implementation Science? *PET Clin*. 2021;16:643-653.

32. Martinho A, Kroesen M, Chorus C. A healthy debate: Exploring the views of medical doctors on the ethics of artificial intelligence. *Artif Intell Med*. 2021;121:102190.

33. Hasani N, Morris MA, Rhamim A, et al. Trustworthy Artificial Intelligence in Medical Imaging. *PET Clin*. 2022;17:1-12.

34. Kilbride MK, Joffe S. The New Age of Patient Autonomy: Implications for the Patient-Physician Relationship. *JAMA*. 2018;320:1973-1974.

35. Lyell D, Coiera E. Automation bias and verification complexity: a systematic review. *J Am Med Inform Assoc*. 2017;24:423-431.

36. Vinuesa R, Azizpour H, Leite I, et al. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat Commun*. 2020;11:233.

37. Jha AK, Myers KJ, Obuchowski NA, et al. Objective Task-Based Evaluation of Artificial Intelligence-Based Medical Imaging Methods:: Framework, Strategies, and Role of the Physician. *PET Clin*. 2021;16:493-511.

38. Grant ES. Requirements engineering for safety critical systems: An approach for avionic systems. In: 2016 2nd IEEE International Conference on Computer and Communications (ICCC). ; 2016:991-995.

39. Zhou Q, Zuley M, Guo Y, et al. A machine and human reader study on AI diagnosis model safety under attacks of adversarial images. *Nat Commun*. 2021;12:7281.

40. Hasani N, Farhadi F, Morris MA, et al. Artificial Intelligence in Medical Imaging and its Impact on the Rare Disease Community: Threats, Challenges and Opportunities. *PET Clin*. 2022;17:13-29.

41. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*. 2019;1:206-215.

42. Harvey DL. Agency and community: A critical realist paradigm. *J Theory Soc Behav*. 2002;32:163-194.

43. Science and Technology Policy Office. Notice of Request for Information (RFI) on Public and Private Sector Uses of Biometric Technologies. *Federal Register*. 2021;86:56300-56302.

44. Saboury B, Rahmim A, Siegel E. PET and AI Trajectories Finally Coming into Alignment. *PET Clin*. 2021;16:xv-xvi.

45. Saboury B, Rahmim A, Siegel E. Taming the Complexity: Using Artificial Intelligence in a Cross-Disciplinary Innovative Platform to Redefine Molecular Imaging and Radiopharmaceutical Therapy. *PET Clin*. 2022;17:xvii-xix.

46. Reader AJ, Schramm G. Artificial Intelligence for PET Image Reconstruction. *J Nucl Med*. 2021;62:1330-1333.

47. Shiri I, Ghafarian P, Geramifar P, et al. Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC). *Eur Radiol*. 2019;29:6867-6879.

48. Yu Z, Rahman MA, Schindler T, Laforest R, Jha AK. A physics and learning-based transmission-less attenuation compensation method for SPECT. *Proc SPIE Int Soc Opt Eng*. 2021;11595.

49. Shiri I, Arabi H, Geramifar P, et al. Deep-JASC: joint attenuation and scatter correction in whole-body 18F-FDG PET using a deep residual network. *Eur J Nucl Med Mol Imaging*. 2020;47:2533-2548.

50. Liu F, Jang H, Kijowski R, Zhao G, Bradshaw T, McMillan AB. A deep learning approach for 18F-FDG PET attenuation correction. *EJNMMI Phys*. 2018;5:24.

51. Van Hemmen H, Massa H, Hurley S, Cho S, Bradshaw T, McMillan A. A deep learning-based approach for direct whole-body PET attenuation correction. *J Nucl Med*. 2019;60:569-569.

52. Rahman A, Zhu Y, Clarkson E, Kupinski MA, Frey EC, Jha AK. Fisher information analysis of list-mode SPECT emission data for joint estimation of activity and attenuation distribution. *Inverse Probl*. 2020;36.

53. Qian H, Rui X, Ahn S. Deep Learning Models for PET Scatter Estimations. In: 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). ieeexplore.ieee.org; 2017:1-5.

54. Arabi H, Bortolin K, Ginovart N, Garibotto V, Zaidi H. Deep learning-guided joint attenuation and scatter correction in multitracer neuroimaging studies. *Hum Brain Mapp*. 2020;41:3667-3679.

55. Sanaat A, Arabi H, Mainta I, Garibotto V, Zaidi H. Projection Space Implementation of Deep Learning-Guided Low-Dose Brain PET Imaging Improves Performance over Implementation in Image Space. *J Nucl Med*. 2020;61:1388-1396.

56. Yu Z, Rahman MA, Schindler T, et al. AI-based methods for nuclear-medicine imaging: Need for objective task-specific evaluation. *J Nucl Med*. 2020;61:575-575.

57. Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. Deep learning in medical image registration: a review. *Phys Med Biol*. 2020;65:20TR01.

58. Yousefirizi F, Pierre Decazes, Amyar A, Ruan S, Saboury B, Rahmim A. AI-Based Detection, Classification and Prediction/Prognosis in Medical Imaging:: Towards Radiophenomics. *PET Clin*. 2022;17:183-212.

59. Cui J, Gong K, Guo N, Kim K, Liu H. CT-guided PET parametric image reconstruction using deep neural network without prior training data. *Medical Imaging 2019*. 2019.

60. Xie N, Gong K, Guo N, et al. Clinically Translatable Direct Patlak Reconstruction from Dynamic PET with Motion Correction Using Convolutional Neural Network. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Springer International Publishing; 2020:793-802.

61. Gong K, Catana C, Qi J, Li Q. Direct Reconstruction of Linear Parametric Images from Dynamic PET Using Nonlocal Deep Image Prior. *IEEE Trans Med Imaging*. October 2021.

62. Jackson P, Hardcastle N, Dawe N, Kron T, Hofman MS, Hicks RJ. Deep Learning Renal

Segmentation for Fully Automated Radiation Dose Estimation in Unsealed Source Therapy. *Front Oncol*. 2018;8:215.

63. Akhavanallaf A, Shiri I, Arabi H, Zaidi H. Whole-body voxel-based internal dosimetry using deep learning. *Eur J Nucl Med Mol Imaging*. 2021;48:670-682.

64. Langlotz CP, Allen B, Erickson BJ, et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology*. 2019;291:781-791.

65. Morris MA, Saboury B, Burkett B, Gao J, Siegel EL. Reinventing Radiology: Big Data and the Future of Medical Imaging. *J Thorac Imaging*. 2018;33:4-16.

66. Sitek A, Ahn S, Asma E, et al. Artificial Intelligence in PET: An Industry Perspective. *PET Clin*. 2021;16:483-492.

67. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25.

68. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*. 2020;580:252-256.

69. Huang S-C, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *NPJ Digit Med*. 2020;3:136.

70. Kaissis G, Ziller A, Passerat-Palmbach J, et al. End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nature Machine Intelligence*. 2021;3:473-484.

71. Warnat-Herresthal S, Schultze H, Shastry KL, et al. Swarm Learning for decentralized and confidential clinical machine learning. *Nature*. 2021;594:265-270.

72. Begoli E, Bhattacharya T, Kusnezov D. The need for uncertainty quantification in machine-assisted medical decision making. *Nature Machine Intelligence*. 2019;1:20-23.

73. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2:158-164.

74. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health*. 2021;3:e745-e750.

75. Arun N, Gaw N, Singh P, et al. Assessing the Trustworthiness of Saliency Maps for Localizing Abnormalities in Medical Imaging. *Radiol Artif Intell*. 2021;3:e200267.

76. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366:447-453.

77. Murray E, Treweek S, Pope C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med*. 2010;8:63.

78. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. *J R Soc Med*. 2011;104:510-520.

79. May C. A rational model for assessing and evaluating complex interventions in health care. *BMC Health Serv Res*. 2006;6:86.

SUPPLEMENTAL INFORMATION

OPPORTUNITIES

In what follows, we discuss a number of opportunities for AI towards improved image generation (part A), image analysis (part B), and radiopharmaceutical therapies (RPTs) (part C). We also refer the reader to a special AI issue of *PET clinics* (*44,45*) edited by and with contributions from a number of SNMMI AI Task force members, with different chapters elaborating a number of the below-mentioned dimensions.

Part A: Opportunities For AI Toward Improved Image Generation

Image Reconstruction. Deep learning models, either standalone or as part of a traditional reconstruction framework, can lead to significant improvements in image quality while achieving reductions of injected activities and/or scan durations. Recent review papers present detailed discussions of different deep learning approaches for image reconstruction (5,46). Standalone AI-based image reconstruction models may also learn the imaging physics of mapping projection data to images. These models are expected to require large volumes of training data comprising paired sets of perfect “ground truth” reconstructed images and raw data, and their development is an opportunity ripe for exploration.

Hybrid approaches may be a practical area of concerted effort, combining neural networks with traditional approaches. In ‘unrolled iterative’ reconstruction methods, a traditional reconstruction framework may be combined with a neural network--for instance an artifact noise reduction model run inside the iterative reconstruction loop. These methods are able to take advantage of known imaging physics, statistics, and data corrections to optimize the reconstruction. AI models can also be utilized to learn the best regularization parameters that have been traditionally challenging to optimize in Bayesian (regularized) image reconstruction methods.

Data Corrections (Attenuation, Scatter, Motion, Denoising). A number of AI-based attenuation and scatter correction methods have been reviewed elsewhere (*6*). Attenuation correction (AC) in SPECT and PET is a prerequisite for quantification and has been shown to be beneficial for visual interpretation tasks. However, AC requires an attenuation map, typically obtained from a CT scan. AI-based methods are promising for AC without requiring CT scans (*28,47*) and are demonstrating promise in both detection (*48*) and quantification (*49*) applications. AI based methods have shown significant potential to discover patterns in sinogram data or images that enable compensation for attenuation and scatter. For example, CNNs can learn quantitative and spatial associations between features in uncorrected PET images and features in paired CT images, allowing algorithms to predict CT images directly from uncorrected PET images (*50*). Or, data corrections can be directly applied without an intermediate step of CT generation (*51*).

AI-based AC can help avoid the increased radiation dose associated with CT-based AC and quantitation errors due to misalignment between the CT image and the PET or SPECT image. This is of particular relevance in multi-time-point acquisitions. AI-based AC is particularly advantageous for PET/MR systems, which lack in-built CT scanning capabilities as well as standalone PET or SPECT systems (*52*).

Deep learning models could be used for generating total scatter profiles (including single and multiple scatters) directly from emission and attenuation sinograms (*53*). The problems of scatter correction and attenuation correction can even be tackled jointly using AI by training end-to-end models that map uncorrected input emission data to attenuation and scatter corrected outputs, early results for which have been shown to be promising in PET (*54*). Yet another promising emerging area of research involves AI models for generating displacement fields required to perform motion-compensated image reconstruction for organs affected by respiratory or cardiac motion.

AI methods are also promising for improving the resolution and noise characteristics of raw sinogram datasets. In particular, AI models have been used towards sinogram denoising (*55*).

Post-Reconstruction Image Enhancement. Recent review papers have discussed the significant opportunities to use AI for further enhancing reconstructed nuclear medicine images (7). Typically, most AI models for post-reconstruction enhancement of images seek to achieve noise reduction or resolution recovery. Unlike image reconstruction approaches, which require access to raw sinogram or list-mode data, post-reconstruction enhancement only uses images for training and validation. Images are more readily available to most users than raw data. Additionally, there are growing public repositories that store medical image data but access to raw data is limited. As a result, AI models for post-reconstruction enhancement are more practical to implement for a larger base of users than AI-based image reconstruction models. Initial attempts for AI-based image denoising and deblurring in this field employed supervised learning models, that rely on matched pairs of corrupt (i.e., noisy and/or low-resolution) and clean (i.e., low-noise/noiseless for the denoising problem and high-resolution for the deblurring problem) images for model training. Paired clinical datasets are not readily available thereby limiting the overall utility of supervised image enhancement techniques, despite their high best-case performance. As a result, a majority of recent efforts in this area have focused on developing unsupervised techniques that use only corrupt images for training and weakly supervised techniques that use unpaired cleaned/corrupt images from separate cohorts. While these image-enhancement methods present strong promise, a key requirement for clinical application will be evaluation on clinical tasks. Current evaluation strategies often use fidelity-based metrics (such as structural similarity index), but studies are showing that this may not correlate with performance on clinical tasks (56). Thus, observer-based studies that focus on task performance are recommended.

Part B: Opportunities For AI Towards Improved Image Analysis

Multiple-Study Image Alignment. It is expected that temporal analysis of imaging changes over time will become increasingly important in the era of enhanced image analysis. In order to compare information from multiple types of imaging examinations (both structural and functional) and multiple time points, a first critical step is appropriate alignment of the images with each other. During different exams the patient may be positioned slightly differently on the examination table. Physiological motion and changes in position of the internal organs particularly the heart, diaphragm, and bowel also occur. Improving the efficiency and reliability of registration is an area of opportunity for AI. Specifically, multiple region-of-interest registration and mass preserving deformable registration are areas that have proved challenging with existing methods. Neural networks have also shown promise in rapidly evaluating registration performance, which may be helpful in quality control in the clinic, even with more traditional and existing registration methods (*57*).

Organ and Lesion Delineations. Rapid and accurate selection of a specific lesion or organ within an imaging volume will be increasingly important to clinicians at the point of care in order to perform more advanced quantifications or to better follow a patient's disease over time. There is an opportunity for AI to make the process of organ and lesion selection faster and more accurate. Neural networks may provide more flexibility and accuracy compared to threshold approaches, for instance through incorporating user input on accuracy of the selection over time. It is expected that numerous AI solutions will emerge that aim to make routine organ and lesion delineation more readily available in the clinic in order to empower more advanced image analytics (*58*). As an example, routine quantification of metabolic tumor volume (MTV) from FDG PET/CT images, known to be superior to SUV quantification for a number of predictive/prognostic tasks for certain cancers, is expected to be routinely enabled by such methods (e.g. for lymphoma patients). These methods can also be extended to quantification of molecular tumor volume from other radiopharmaceuticals.

Imaging Biomarkers and Beyond. Over the last two decades, the extraction of image biomarkers from nuclear medicine images has been the focus of numerous studies. One of the simplest metrics has been the SUV or change in SUV with treatment as formalized by the PERCIST 1.0 metrics. The imaging biomarker field, is broadly a part of the domain known today as "radiomics", and relied until recently on several sequential steps once images are reconstructed and collected for analysis: lesions, tumours or organs detection and segmentation (the contours in 2D or volumes in 3D are determined), followed by characterization (a number of handcrafted features are calculated to describe the segmented volume), in order to finally be exploited in a modeling step. This modeling step consists of selecting a subset of features for their relevance regarding the task at hand (e.g. predicting outcome or differentiating between benign and malignant lesions) and combining them into a multiparametric model. Despite numerous studies and promising results, advanced radiomics has not been widely translated to clinical practice due to several intrinsic limitations: (i) lack of automation (especially for the detection and segmentation of lesions), (ii) lack of standardization (which now has been significantly addressed by the Image Biomarker Standardization Initiative, IBSI, with further ongoing efforts), (iii) harmonization issues (some handcrafted features are notoriously sensitive to numerous factors including scanner device characteristics and performance, reconstruction algorithms and settings and acquisition protocols, which makes radiomic models perform poorly in a real heterogenous setting such as for example multicenter studies), and (iv) explainability and interpretability issues (radiomic features are often quite unintelligible for end users, so models combining several of them can be seen as an untrustworthy "black box").

Over the last few years, there have been several developments that have addressed these issues. The efforts by the IBSI to standardize both features and the workflow have made published studies more comparable. The rise of techniques based on deep neural networks has already provided solutions for the lack of automation (e.g. methods based on the U-Net architecture are now providing more effective, automated detection and delineation of lesions in PET images) and may also improve modeling and harmonization issues, provided appropriate training strategies are implemented and large enough datasets for training are made available. Given their impressive performance obtained in various computer vision applications, it is expected that the usual radiomics workflow currently implemented as sequential steps and the extraction of handcrafted features from delineated lesions might end up being replaced by end-to-end approaches relying on deep neural networks using as input not the delineated tumor volume but simply the entire PET image to learn relevant features contained in the images and derive its output. Though it is possible that these networks may be expanded to additionally incorporate certain handcrafted features (e.g. shape features) that may not be easily captured by neural networks with limited training. Overall, "handcrafted" and "deep" radiomics are areas of significant ongoing activity.

Kinetic Modeling. Quantitative and qualitative improvements in reconstructed dynamic images directly translate to enhanced accuracy in the spatial maps of kinetic micro- and macro-parameters (59). Particularly, because of the characteristics of tracer dynamics, there are commonly some short static frames in a dynamic PET scan. These short frames lead to noisy PET images which pose a critical challenge to robust kinetic modeling and parametric image computation. With the superior performance of deep learning-based image reconstruction, particularly for low count rates (i.e. short frames), the resulting parametric images can have enhanced accuracy. In addition, with the flexibility of deep neural networks to map complex functions, one can synthesize high quality parametric images from low quality parametric images (60), or directly reconstruct parametric images from raw data (61) leading to even better image quality. Application of AI in the context of kinetic modeling as applied to large-axial-FOV or total-body PET is another exciting frontier (18).

Part C: Opportunities For AI Toward Improved Radiopharmaceutical Therapies

Personalized Dosimetry. There are multiple steps within dosimetry as applied to RPTs that will be readily enabled and enhanced for routine deployment via AI methods. To integrate multiple time-point imaging data, registration of images, as well as segmentation of organs and tumors is required. These processes can be time-consuming, cumbersome, and pose a practical-limitation to routine dosimetry. AI might improve the accuracy of whole-body multiple time-point image registration (*57*), and has been shown to be capable of automatically segmenting organs at risk and target tumors (*62*).

Following multi time-point SPECT acquisition, activity concentrations of various organs and tissues are plotted over time to obtain time activity curves (TAC). Due to the limited number of time-points available, these TACs are fit to a curve, extrapolating beyond the final time-point, and then integrated to yield the cumulated activity, or time integrated activity (TIA). AI based curve fitting could eventually reduce the number of time points needed (*25*).

Converting the TIA map (macro-level radiopharmaceutical gamma signature) to tissue-level effective dose of the particles (electron or alpha particle) can involve complex modeling of particle trajectory, its interaction with matter, and micro-level sensitivity of the biological tissue. This could be an opportunity for AI because it is currently extremely cumbersome and computationally intensive (*63*).

The cumulative effect of each treatment session is a function of dose deposition in each session, time-interval between the sessions, and other biological conditions of the patient (tissue sensitivity, repair capacity, immune system response, etc). There is an important potential role for AI in the creation of this comprehensive patient dose profile and making it practically and ubiquitously available. This profile could aid the nuclear medicine physician in improved adaptive treatment planning, through optimal time-interval determination and dose prescription, as well as utilization of adjunct measures such as chemotherapy and immunotherapy. Specifically, post-treatment dose deposition quantification in target lesions will ensure sufficient therapy and identify the under-treated subsections for “proactive mitigations”, such as locoregional therapies (interventional oncology and radiation oncology) as well as augmented therapy (adjuvant RPT or chemotherapy). Post-treatment dose deposition quantification in organs at risk (OAR) could guide optimal treatment planning while minimizing normal organ damage. Such approaches can be especially enhanced through utilizing digital twins, as described next.

Digital Twins. A digital twin is the numeric representation of a patient spanning his or her entire life. The patient’s digital twin can be (i) updated with real-time data (e.g. cumulative radiation dose) reflecting history and current condition of the patient, and (ii) can be used for simulations to aid physicians in complex treatment planning scenarios. This concept, first coined by Michael Grieves, has a proven track record of success in modeling complex industrial engineering applications. Radiopharmaceutical therapies (RPTs) have key factors that distinguish their planning from external radiation therapy and increase their complexity. Examples include distinct pharmacokinetics, dose rates, temporal scales (including temporal heterogeneity), spatial scales (including spatial heterogeneity), and linear energy transfer (LET) rates.

The theranostic digital twin (TDT) incorporates a combination of structural imaging and dynamic molecular imaging to produce a pharmacokinetic biodistribution model that is specific to a particular patient—quantifying normal organ systems biology and tumor biology modeling. The result is the ability to model normal tissue complication probability and tumor control probability (TCP) based on the modeled biological effective dose. TCP computation is a dynamic interplay between dose rate pharmacokinetics, cell type, genomics, DNA repair pathways and other factors such as hypoxia. The TDT coupled with appropriate computational tools can be used for predictive dose modeling; e.g. a model can be personalized based on pre- and/or intra-therapy molecular imaging. Different injection strategies and intervals can be explored. Mitigation strategies for suboptimal dose delivery can also be pursued such as adaptive dose planning, augmentation with locoregional therapy (e.g. ablative therapy or external radiation therapy), or adjuvant systemic strategies such as chemotherapy, immunotherapy, and CAR-T therapy.

Each time a patient receives an RPT, the TDT can be updated to record the acquired information related to his or her healthy and diseased tissues. This data could be stored locally, on a patient's personal device, or in a cloud resource such as a healthcare exchange platform so that it could be dynamically updated as the patient receives subsequent procedures. In addition to practical considerations such as nuclear medicine treatment and re-treatment planning, the TDT could be used over time to improve subsequent image reconstructions. Prior image labeling tasks can also be catalogued by the digital twin to optimize subsequent follow up exam evaluation and interpretation efficiency by nuclear medicine physicians.

The TDT may greatly improve the reliability of quantification measures compared to the current standard of practice, particularly radiation dose over time. The digital twin is a next evolution in personalizing medical imaging and will be aided by AI: as outlined earlier, AI can significantly assist in enabling rapid and reliable dosimetry calculations; further the complex biological models can be replicated (with sufficient training) using appropriate neural networks.

CHALLENGES FOR DEVELOPMENT, VALIDATION, DEPLOYMENT, AND IMPLEMENTATION

On balance with the potential opportunities for AI to impact the field of nuclear medicine, there are also obstacles. In this section, we review current challenges effecting development, validation, regulation, dissemination/implementation, and public trust.

Development Challenges

Data. One of the greatest impediments to AI research and development in medical imaging is the availability of data, which should be findable, accessible, interoperable, and reusable (FAIR) (64). Current publicly available data is of limited volume, from relatively few institutions, with a relatively narrow range of disease representations. At the time of this article, there are only 2346 publicly available PET imaging subjects in the cancer imaging archive (TCIA), which cover predominantly head and neck, lung, and breast cancer. There is also a need for an increased ability for patients to easily and securely share their medical data and to assure appropriate patient privacy protections. However, having access to imaging and clinical data is not enough. There is no readily available process of converting big data (data lakes) into more organized datasets (data warehouses) in medicine (65). In order to be useful for the development of AI applications, clinical and imaging data must be curated to include standardized descriptors (metadata—including highly clinically-relevant data such as survival), labels, and must be appropriately deidentified (AI readiness).

Data should be able to be compared and evaluated across patient populations, diseases institutions, and among various vendors. Raw imaging data, which is richer than reconstructed data, is almost always purged on a regular basis due to its sheer size requirements and proprietary nature, even in modern clinical and research archives. There is no universal format for list-mode data (66). Quality and variability of input data are critically important and may affect outcomes, so data sets of reasonably uniform quality are likely of great value, though it can be argued that the most robust AI methods would function well on a diverse range of data inputs of differing qualities.

Optimal Network Architecture. Many published works on artificial intelligence in medical imaging employ transfer learning from CNNs trained on non-medical 2-dimensional imaging data(67). Dedicated medical imaging datasets are potentially preferable, although current available sets remain small in comparison to the non-medical imaging datasets. This limitation was highlighted as a key challenge in the 2018 National Institutes of Health roadmap (64).

Much of nuclear medicine data have 3 dimensions or more (e.g. x, y, z and sometimes time). AI architecture designed to incorporate additional dimensions (such as those designed for video recognition tasks has shown promise toward incorporating contextual information in other areas. However, there is limited literature available and video CNN training can be orders of magnitude more computationally challenging (68). Analyses including registered CT and MRI data with nuclear medicine may also prove informative.

More complex architectures will also need to be considered. Fusion architecture may also be needed, such as paired data to incorporate deep natural language processing CNNs with computer vision CNNs for certain tasks (69). Federated and swarm learning approaches in medical imaging that are privacy-preserving and secure while addressing issues of network latencies will be needed to train large-scale task-specific datasets (70,71).

Measurement and Communication of Uncertainty. There is a need for estimation and reporting of uncertainty with each AI system output in medicine. Classifications often provide an output category without conveying the uncertainty of that assignment. In cases where the estimation of uncertainty is conveyed, there are minimal assurances that the estimation itself is accurate. Without these metrics, automation bias risk is left unchecked (35). AI applications in nuclear medicine must embrace these principles, which have proved helpful to other areas (72).

Clinically Impactful Use Cases. Developers of AI need help to understand the most clinically important needs where AI could provide added utility. There is currently a mismatch between the use cases explored by many researchers or vendors and the clinical needs. This is, to a large degree, related to the limited availability of public datasets. A forum for improved direction including the most viable areas to expend development resources informed by expert professionals in the field of nuclear medicine does not yet exist.

Team Science. There is limited awareness of the potential for AI in nuclear medicine by much of the computer science development community. Nuclear medicine AI developers need engagement opportunities (37). Collaboration among scientists in many domains is needed in order to realize the scope of potential. An environment of competition to serve as a launchpad for attracting talent and stimulating interdisciplinary approaches is yet to be established in the field of nuclear medicine AI.

Evaluation Challenges

Performance Profiling Through Task-Based Evaluations. We need to understand when, and if possible, why AI software fails (*failure mode profiling*). Although AI models may help clinicians to identify *novel features* of clinical importance (73), AI models can also make decisions based on *irrelevant or non-specific features*. While in some instances failure mode profiling could be aided by AI explainability, tackling the explainability hurdle alone is not sufficient (74). Controversy remains in the literature for saliency maps (popular methods highlighting 'relevant' portions of images for specific tasks) as to how effective and beneficial they are (75) or even whether understanding CNN failure modes is practical in high stakes decisions (41). One approach to address this issue is rigorous evaluation on populations of patients and on clinical tasks. The task-based evaluation paradigm (population-based or personalized) provides for an approach to address these challenges; however strategies to conduct such evaluation are still being explored (*4*,37). Population-based evaluation must be further augmented by investigating incidents when failure occurs specific to an individual patient (personalized evaluation).

Bias and discrimination are also important issues difficult to evaluate. Approaches to evaluate AI software for known or unknown biases are important (*76*).

Guidelines for Validation. There are few guidelines for appropriate initial and continuous validation of AI software in medicine. There is a need for a framework for the evaluation of AI software designed for use in nuclear medicine in the real world. A proof-of-concept should be provided including an objective, rationale, study design, and output measures. Well-defined testing procedures and reference standards must be available for the tasks at hand. It must be realized that a balance exists between moving AI approaches forward in a timely manner and their perfection in broad populations and a wide range of clinical use cases.

Multi-Center Clinical Trial Network. There is a need for a clinical trial network composed of multiple institutions to cross-validate AI software among different clinical environments.

Ethical, Regulatory, and Legal Ambiguities

Ethical Aspects. Ethical issues related to AI systems in nuclear medicine are only beginning to be defined.

Although the classic principles of medical ethics (autonomy, non-maleficence, beneficence, and justice) are well established, the complexity of modern clinical decision-making requires reconciliation of conflicting principles. Ethical dilemmas could be more challenging in the presence of AI devices in healthcare. One reason for this is due to the role of patient data in the genesis of new AI software and the complexities of informed consent. Another issue is that insinuated replication of historical biases and unfairness embedded in training data could be imparted in the produced AI device. There are also issues related to the appropriate agency of the AI device, potential unintended consequences it presents, and the inherent opaqueness of some of the useful algorithms.

Furthermore, the scope of ethical dilemmas is broader than the traditional dyadic relation between physician and patients. There are more stakeholders sharing in both the benefits and the burdens. There is a need for inclusion and active participation of all stakeholders in order to understand how to best resolve ethical questions.

Finally, there are legitimate concerns about the impact of these technologies on healthcare disparities on one hand and their trustworthiness on the other. There are no established guidelines for the appropriate inclusivity of AI systems in terms of how to address bias, discrimination, or other issues that may arise. Although limitations may be unavoidable in certain approaches, we must better understand what are acceptable limitations and what are appropriate compensatory mitigations when necessary. In reality, access to high level nuclear medicine patient care varies across geographic areas with rural and some central urban areas underserved medically. The availability of AI tools may help these populations to a greater extent than the benefits accrued to populations more richly served with specialists. Thus, delaying deployment of AI tools until they are "perfected" may limit benefits generated by such algorithms.

The intricacies of ethical AI in medicine leaves many questions that remain to be solved. We should recognize the necessity of more comprehensive ethical discourse, such as collectively deliberated contracts to respect equal basic rights, ensure fair sharing of benefits and burdens, and emphasize the importance of fair deliberative processes.

Regulatory and Legal Aspects. There is limited awareness by the public and stakeholders about the level of evidence required to validate safety, security (39) and each specific clinical claim by AI software solutions, as well as the evidential requirements to verify the added-value for appropriate reimbursement.

There is limited legal precedent for the use or misuse of AI in healthcare and it is even unclear whether product liability law would apply to AI software in medicine, particularly if the software changes over time after regulatory approval (29). At present, we believe that all AI in nuclear medicine must be supervised by a physician and that patient care choices can be informed by AI, but ultimately are made by the physician based on all data and the doctor/patient relationship. Thus, we are not currently proposing that "autonomous AI" would eliminate the physician from patient care, but rather that the AI would augment physician decision making.

Implementation of Clinical AI Solutions & Post-Deployment Monitoring

AI-Platform. Current platforms for the integration of AI software applications into the clinical workflow are cumbersome. AI software must be able to work within the clinical context. Appropriate platforms are needed that allow nuclear medicine physicians to select and utilize various AI tools independent of a particular vendor—e.g. to have the AI modules as "plug-ins" or apps complementing PACS or NM specific display systems (30).

Barriers of Dissemination and Implementation of AI Technology in Medicine. Normalization of health interventions is the collective action to incorporate new changes into everyday practice so they become institutionalized and disappear from view (thus normalized) (77). According to *Normalization Process Theory*, the process of implementation involves various interconnected steps at the individual, institutional and societal levels. Adoption of any new evidence in the healthcare system takes considerable time, on average 17 years (78). With regards to the specific adoption of AI technology, the process of dissemination and implementation might be even more complicated (31).

At the individual level, one of the major challenges is how healthcare providers perceive the utility of AI software. Is it helpful or another 'hype'? Does AI save time or decrease throughput and overall efficiency? This perception could be held regardless of evidence-based performance metrics. In addition, healthcare providers must possess the skills to use it or be willing and able to attain these skills. The system needs workability within the context of the clinical workflow (79). There must be an incentive to change practice patterns to incorporate the new solution. The AI system as a new agent in the ecosystem could be perceived as a rival or a partner.

At the institutional level, the adoption of AI can pose new challenges which must be weighed against other institutional needs. Institutional culture can be in favor or against the change. The commitment to incorporate new change (cognitive participation) is very important and challenging. This is essential for sustainable group coherence to collectively engage in the act of implementing (collective action), and continuous evaluation and revision (reflexive monitoring).

At the societal level, there are multiple factors including norms, social roles, regulations and oversights, reimbursement policies, legal frameworks, and public material and informational resources that impact dissemination and implementation.

Post-Deployment: Change Management & Performance Monitoring. AI software may change over time as new data are integrated into a model. The total product life cycle approach must therefore be considered. Change control methodology must be established in a way that supports improvement and protects patient safety. There is a need for methods to be able to conduct post-market surveillance for AI software in medicine by regulatory bodies, much as there is for current medical software.

STRATEGIES FOR SUCCESS

SNMMI Initiatives

TO ADDRESS the need for guidance toward best AI system practice, SNMMI created the AI Task Force, aiming to monitor and explore emerging issues in the field of AI, to identify opportunities and challenges, and to recommend appropriate actions, policies, and programs to the society's governing bodies and members. The task force comprises 4 teams, focusing on (i) strategic planning, partnership, and outreach (SPO), (ii) algorithm development, (iii) evaluation, and (iv) ethical considerations. This paper is the report of those deliberations.

TO ADDRESS the need for inter-disciplinary creative collaboration for AI development, SNMMI AI Task Force designated the "SPO working group" to plan, organize and host the

SNMMI AI Summit to bring together subject matter experts from academia, industry, non-governmental organizations, and government agencies, as well as physicians, physicists, scientists, technologists, and other stakeholders toward the continuous improvement of the *AI and Informatics Ecosystem* for nuclear medicine.

TO ADDRESS the need for best practices for algorithm development, SNMMI AI Task Force designated the "Development working group" to prepare recommendations and guidelines. Challenges and pitfalls to AI algorithm development were identified, and appropriate methods for study design, data collection and curation, algorithm development and testing, and reporting/dissemination were proposed. Additional recommendations specific to various subspecialties within nuclear medicine were also provided (*3*).

TO ADDRESS the need for appropriate evaluation of AI systems, SNMMI AI Task Force designated the "Evaluation working group" to prepare recommendations and guidelines. The key factors to consider in an evaluation study were envisioned, including the need to assess generalizability and performance on clinical tasks. The working group put forth the view that an evaluation study should result in a claim (*4*). A comprehensive four-class evaluation framework was established consisting of proof-of-concept, technical efficacy, clinical evaluation, and post-market deployment studies. For each class of evaluation, recommendations for data collection, curation, sample-size determination, quantitative metrics to ascertain success, and example claims were provided.

TO ADDRESS the ethical aspects of AI development and implementation, SNMMI AI Task Force designated the "Ethics working group" to contemplate on the topic through the engagement of all of the stakeholders (communicative action).

TO ADDRESS the emerging needs in the realm of precision radiopharmaceutical therapy, SNMMI AI Task Force designated the "AI & Dosimetry working group" to investigate the role of AI in multi-scale dosimetry, predictive dosimetry (treatment planning) and post-treatment dosimetry (treatment verification).

TO ADDRESS the educational needs of the nuclear medicine community, SNMMI AI Task Force designated the "Data Science & AI Curriculum working group" to prepare relevant educational material for four distinct audiences including practicing attending physicians, practicing nuclear medicine physicists, practicing nuclear medicine technologists, and in-training nuclear medicine professionals.

SNMMI Action Plan

Is AI a fundamental element of all aspects of nuclear medicine or an entity of its own?

It can be argued that the current structure of the SNMMI should absorb AI into all of its activities. For example, AI will be an increasing component of the Physics Data and Instrumentation space. Similarly the performance of AI could easily be a part of the quality and evidence committee while regulatory aspects of AI can be part of the governmental affairs committee. However, an alternative approach is to treat AI as a "separate" entity for a period of time. This could be done in the following ways:

TO ADDRESS the perpetual needs for trustworthy AI, SNMMI AI Task Force has recommended the establishment of the AI Center of Excellence (AICE). Each of the working groups formed under the AI task force could become lasting committees of AICE including SPO

Committee, Development Committee, Evaluation Committee, Ethics Committee, and Data Science & AI Curriculum Committee. In addition, AICE is recommended to consider construction of three new entities, as detailed below.

TO ADDRESS the nuances of AI Development, SNMMI AI Task Force recommends AICE Development Committee to consider task specific guidelines such as best practices for development of AI-based image reconstruction or AI-based image segmentation.

TO ADDRESS the intricacies of AI Performance Evaluation, SNMMI AI Task Force recommends that the AICE Evaluation Committee further detail the Task-Based Assessment Framework (*4*). An outline for conducting such evaluation was recently proposed, including strategies to conduct such an evaluation (37).

TO ADDRESS the need for an impartial performance evaluation of AI systems available for clinical use pertaining to the practice of nuclear medicine, SNMMI AI Task Force recommends AICE to establish an AI Clinical Trials Network (AI-CTN). AI-CTN will seek to collaborate with other professional society counterparts and work toward interfacing with industry to design and implement a sustainable-development-ecosystem.

TO ADDRESS the need for nuclear medicine data for the development and evaluation of AI tools, SNMMI AI Task Force recommends AICE to consider identifying resources to allow formation and sustainability of the Nuclear Medicine Imaging Archive (NMIA). NMIA will serve as a resource for the AICE SPO Committee to host competitive AI challenges to address clinical needs through novel solutions. NMIA will serve to assist in converting nuclear medicine data lakes into organized data warehouses. These AI-ready resources will aid in research and development of AI solutions in nuclear medicine. In addition, NMIA will allow curation of datasets for evaluation of performance.

TO ADDRESS the need for multi-disciplinary and inclusive discourse to actualize the ethical and trustworthy implementation of AI in medicine and society, SNMMI AI Task Force recommends AICE to spearhead formation of the Trustworthy AI in Medicine and Society Coalition (TAIMS coalition to 'tame' AI) with the help of the AICE SPO Committee. This coalition shall include medical imaging societies and non-imaging medical groups as well as non-medical societal institutions (toward The AI Bill of Rights (*43*)).

The above needs are important, yet one has to be cognizant of the needs versus the number of volunteers and resources available to best address the important tasks. Establishing a new SNMMI task force on AI implementation may be an appropriate intermediate structure.

SNMMI Recommendations

Integration of AI Algorithms into Clinical Workflow (AI Orchestrator for Interoperability). To facilitate dissemination and implementation of AI-based algorithms in the clinical setting, workflow integration is needed. Integration of these applications into the clinical image viewer (PACS) must be seamless and 'vendor neutral' in order to be widely adopted. An 'AI Orchestrator' could interface with functionality currently available in a PACS both locally and through cloud applications. The AI Orchestrator would enable physicians and medical providers to select the best-of-breed AI applications without being tied to a particular PACS vendor platform. To actualize this goal, there should be a collaborative effort among AI developers, enterprise imaging archive vendors, interoperability standardization organizations, and professional medical societies.

This recommendation is essential for realization of clinical image interpretation and quantification of tomorrow in which physicians will be able to choose the best and most cost-effective AI technologies for specific clinical indications and will be freer to custom tailor their own workflow as new, better, or more cost-effective solutions become available. With modular components that scale with needs, smaller institutions and individual physicians will more easily incorporate technologies currently only available at much larger healthcare systems.

Harmonized List-Mode Data Format. As demonstrated in other research fields driven by open science, a common data standard, ethical principles and public datasets are the keys to initiate a successful new wave of productive research and growth. In nuclear medicine and molecular imaging, the major vendors have different list mode formats, particularly for time-of-flight information. This is one of the biggest roadblocks for the development of both traditional image formation and analysis algorithms, and becomes a more salient problem in the era of artificial intelligence and open science. To address this point, a vendor neutral list mode format is urgently needed to move the field forward. It is technically not difficult, but needs communication and endorsement among all the shareholders.

Regulatory Process. Regulation of AI software is in the early stages and will continue to evolve. SNMMI recommends professional societies actively engage with each other to share clinical experience of experts practicing in the affected clinical areas, which could be informative to regulatory agencies.

Certification and Accreditation Pathways. Subspecialized tracks should be conceptualized to demonstrate added competency in the clinical aspects of nuclear medicine informatics (26). For board certified nuclear medicine physicians, an ACGME accredited Clinical Informatics fellowship with focus on the nuances of advanced molecular imaging and therapy should be established for subspecialty board certification by American Board of Preventive Medicine (ABPM) in Clinical Informatics.