

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.

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. Saboury B, Rahmim A, Siegel E. PET and AI trajectories finally coming into alignment. *PET Clin*. 2021;16:15–16.
6. 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;17:17–19.
7. Reader AJ, Schramm G. Artificial intelligence for PET image reconstruction. *J Nucl Med*. 2021;62:1330–1333.
8. 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.
9. 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.
10. Shiri I, Arabi H, Geramifar P, et al. Deep-JASC: joint attenuation and scatter correction in whole-body ^{18}F -FDG PET using a deep residual network. *Eur J Nucl Med Mol Imaging*. 2020;47:2533–2548.
11. Liu F, Jang H, Kijowski R, Zhao G, Bradshaw T, McMillan AB. A deep learning approach for ^{18}F -FDG PET attenuation correction. *EJNMMI Phys*. 2018;5:24.

12. 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 [abstract]. *J Nucl Med*. 2019;60(suppl 1):569.
13. 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;36:084002.
14. 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)*. IEEE; 2017:1–5.
15. 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.
16. 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.
17. Yu Z, Rahman MA, Schindler T, et al. AI-based methods for nuclear-medicine imaging: need for objective task-specific evaluation [abstract]. *J Nucl Med*. 2020;61(suppl 1):575.
18. 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.
19. 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.
20. Cui J, Gong K, Guo N, Kim K, Liu H. CT-guided PET parametric image reconstruction using deep neural network without prior training data. In: *Proceedings of SPIE 10948, Medical Imaging 2019: Physics of Medical Imaging*. SPIE; 2019:109480Z.
21. 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.

22. 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*. 2022;41:680–689.
23. 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.
24. 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.
25. 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.
26. 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.
27. Sitek A, Ahn S, Asma E, et al. Artificial intelligence in PET: an industry perspective. *PET Clin*. 2021;16:483–492.
28. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems 25 (NIPS 2012)*. MIT Press; 2012:1–9.
29. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*. 2020;580:252–256.
30. 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.
31. Kaissis G, Ziller A, Passerat-Palmbach J, et al. End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nat Mach Intell*. 2021;3:473–484.

32. Warnat-Herresthal S, Schultze H, Shastry KL, et al. Swarm Learning for decentralized and confidential clinical machine learning. *Nature*. 2021;594:265–270.
33. Begoli E, Bhattacharya T, Kusnezov D. The need for uncertainty quantification in machine-assisted medical decision making. *Nat Mach Intell*. 2019;1:20–23.
34. 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.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. May C. A rational model for assessing and evaluating complex interventions in health care. *BMC Health Serv Res*. 2006;6:86.
41. 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.
42. 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.

43. 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.
44. 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.
45. Cross DJ, Komori S, Minoshima S. Artificial intelligence for brain molecular imaging. *PET Clin.* 2022;17:57–64.
46. Gharavi SMH, Faghihimehr A. Clinical application of artificial intelligence in PET imaging of head and neck cancer. *PET Clin.* 2022;17:65–76.
47. 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.
48. Miller RJH, Singh A, Dey D, Slomka P. Artificial intelligence and cardiac PET/computed tomography imaging. *PET Clin.* 2022;17:85–94.
49. 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.
50. Paravastu SS, Theng EH, Morris MA, et al. Artificial intelligence in vascular-PET: translational and clinical applications. *PET Clin.* 2022;17:95–113.
51. Paravastu SS, Hasani N, Farhadi F, et al. Applications of artificial intelligence in ^{18}F -sodium fluoride positron emission tomography/computed tomography: current state and future directions. *PET Clin.* 2022;17:115–135.
52. 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.

53. 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.
54. 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.
55. Ding W, Yu J, Zheng C, et al. Machine learning-based noninvasive quantification of single-imaging session dual-tracer ^{18}F -FDG and ^{68}Ga -DOTATATE dynamic PET-CT in oncology. *IEEE Trans Med Imaging.* 2022;41:347–359.
56. Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. *Nature.* 2021;596:590–596.
57. Webb EW, Scott PJH. Potential applications of artificial intelligence and machine learning in radiochemistry and radiochemical engineering. *PET Clin.* 2021;16:525–532.
58. Ataieinia 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.
59. 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.
60. 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.
61. 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.
62. Beegle C, Hasani N, Maass-Moreno R, Saboury B, Siegel E. Artificial intelligence and positron emission tomography imaging workflow. *PET Clin.* 2022;17:31–39.

63. Ullah MN, Levin CS. Application of artificial intelligence in PET instrumentation. *PET Clin.* 2022;17:175–182.
64. 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.* 2020;3:e200137.
65. Mezrich JL. Demystifying medico-legal challenges of artificial intelligence applications in molecular imaging and therapy. *PET Clin.* 2022;17:41–49.
66. Saboury B, Morris M, Siegel E. Future directions in artificial intelligence. *Radiol Clin North Am.* 2021;59:1085–1095.
67. 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.
68. 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.
69. Hasani N, Morris MA, Rhamim A, et al. Trustworthy artificial intelligence in medical imaging. *PET Clin.* 2022;17:1–12.
70. Kilbride MK, Joffe S. The new age of patient autonomy: implications for the patient-physician relationship. *JAMA.* 2018;320:1973–1974.
71. Lyell D, Coiera E. Automation bias and verification complexity: a systematic review. *J Am Med Inform Assoc.* 2017;24:423–431.
72. Vinuesa R, Azizpour H, Leite I, et al. The role of artificial intelligence in achieving the sustainable development goals. *Nat Commun.* 2020;11:233.
73. 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.
74. Grant ES. Requirements engineering for safety critical systems: an approach for avionic systems. In: *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*. IEEE; 2016:991–995.

75. 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.
76. 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.
77. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell.* 2019;1:206–215.
78. Harvey DL. Agency and community: a critical realist paradigm. *J Theory Soc Behav.* 2002;32:163–194.
79. Science and Technology Policy Office. Notice of request for information (RFI) on public and private sector uses of biometric technologies. *Fed Regist.* 2021;86:56300–56302.