Breakout Session 3: Track A

Uncertainty Quantification of AI-Based Imaging Algorithms: The Need and Methods

Dr. Abhinav Jha (Moderator)
Assistant Professor, Washington University

Uncertainty Quantification of AI-Based Algorithms for Medical Imaging: The Need and Methods

Supplemental Award title: A framework to quantify and incorporate uncertainty for ethical application of AI-based quantitative imaging in clinical decision making

Abhinav K. Jha, PhD

Department of Biomedical Engineering

Mallinckrodt Institute of Radiology

SNMMI AI Task Force Member

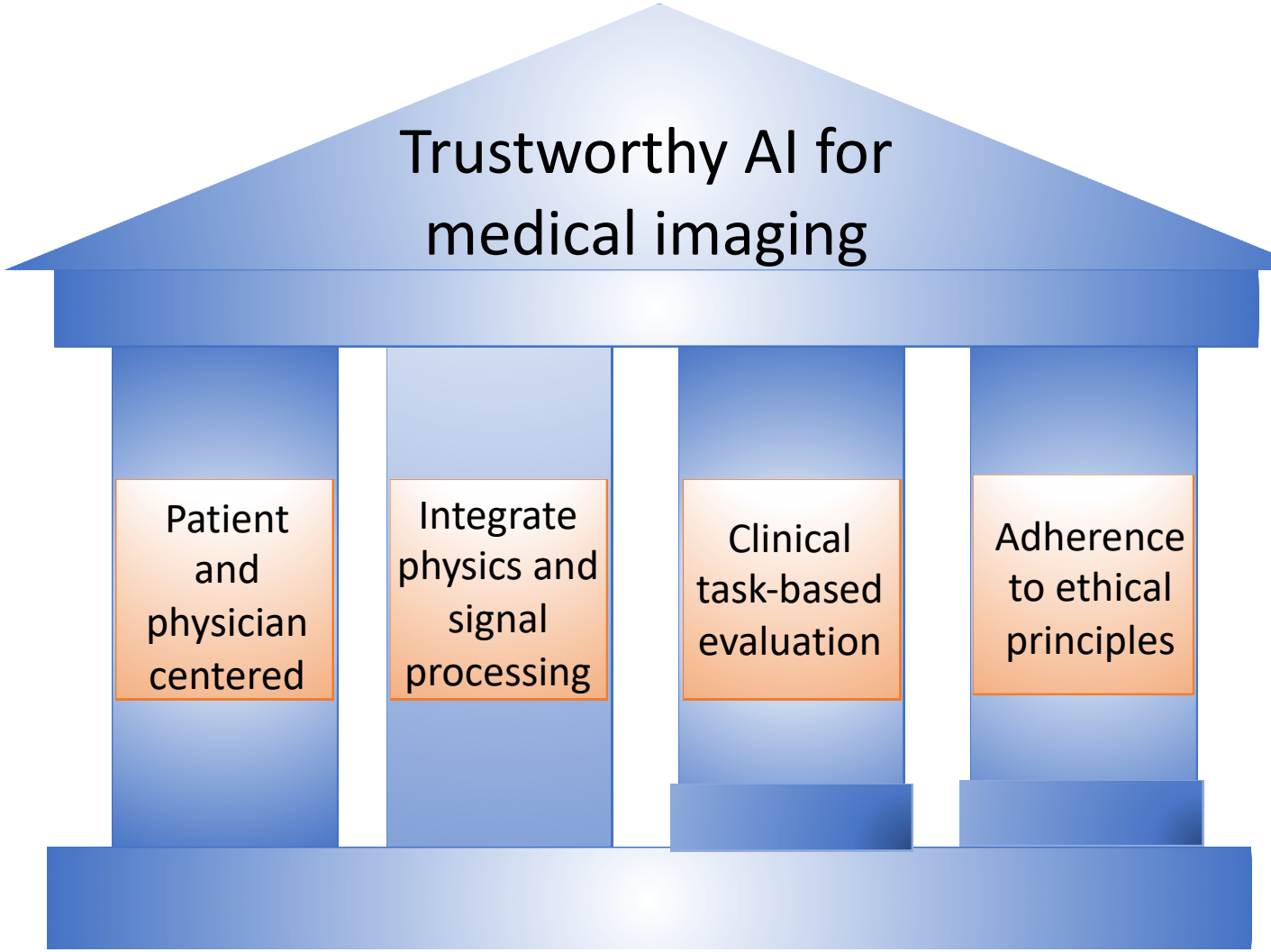
Washington University in St. Louis

McKelvey School of Engineering

School of Medicine, Mallinckrodt Institute of Radiology

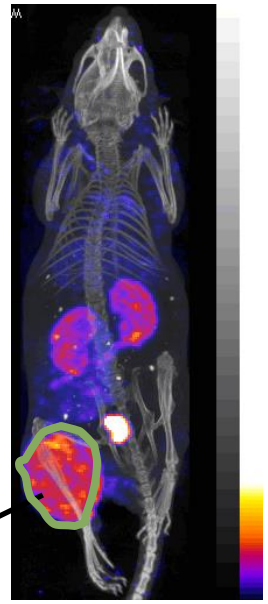
Background

- **Mission of my lab**: Develop computational imaging solutions for improving the diagnosis and treatment of diseases
- **Research expertise**: Image reconstruction, enhancement, quantitative image analysis

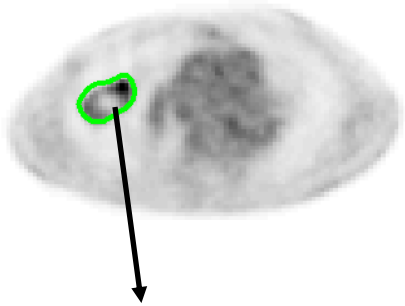


Quantitative medical imaging

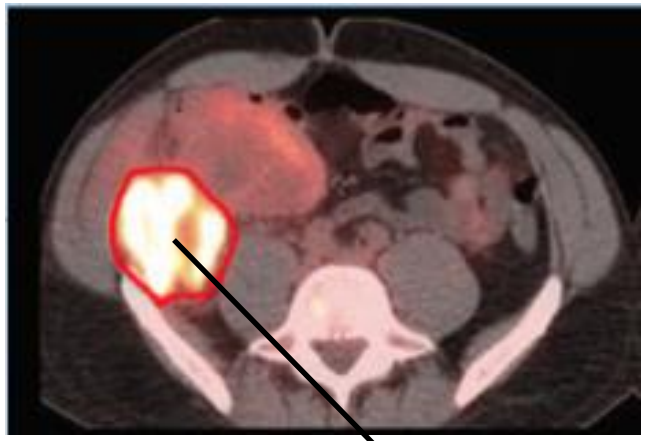
The extraction of numerical/statistical features from medical images for clinical decision making



Tracer uptake



Metabolic Tumor volume (MTV)



Intra-tumor heterogeneity

Quantitative imaging is emerging as an important tool in the clinic for multiple diagnostic and therapeutic procedures

Challenge: Uncertainty in AI algorithm

Anya Plutynski, PhD
Biomedical Ethicist

- AI is playing a strong role in the development of new quantitative imaging methods. However, these algorithms suffer from uncertainty
- Important that uncertainty be accounted for when making clinical decisions
- A patient-advocate survey we conducted indicated likewise
  - ~90% respondents were uncomfortable/very uncomfortable with the idea that uncertainty of AI algorithm is not conveyed to them
  - > 75% respondents specifically wanted uncertainty to be accounted for when making clinical decisions
- Different patients may have different risk-value profiles related to the use of AI in the clinic

Motivation: Modeling uncertainty in AI when making clinical decisions

nature medicine

Explore content ∨ About the journal ∨ Publish with us ∨

nature › nature medicine › comment › article

Comment | Published: 31 January 2022

Clinical decisions using AI must consider patient values

Jonathan Birch ✉, Kathleen A. Creel, Abhinav K. Jha & Anya Plutynski

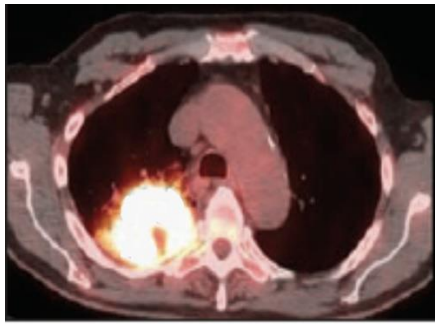
Nature Medicine **28**, 229–232 (2022) | Cite this article

1038 Accesses | **30** Altmetric | Metrics

Built-in decision thresholds for AI diagnostics are ethically problematic, as patients may differ in their attitudes about the risk of false-positive and false-negative results, which will require that clinicians assess patient values.

An important unmet need for strategies to measure and account for uncertainty when using AI-based quantitative imaging for making clinical decisions

Illustration: Project aims



Patient image

AI-based tool to segment tumor

Aim 1
Uncertainty of tumor volume

Oncologist

An informed decision

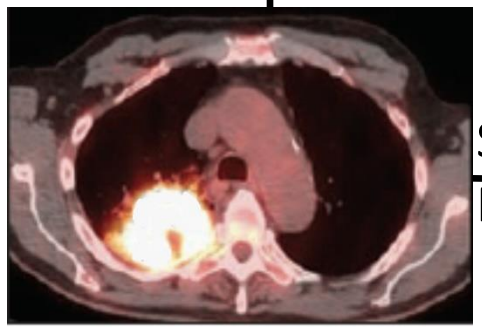
Aim 2
Patient risk-value profile

Patient imaged on scanner

Aim 1: Uncertainty quantification without ground truth

- Existing uncertainty-quantification methods for AI algorithms typically assume the availability of a gold standard.
- Obtaining such a gold standard is time-consuming, expensive, and in many cases, even impossible
- Towards addressing this issue, we have developed an uncertainty quantification technique for evaluating AI-based quantitative imaging methods that does not require ground truth

Uncertainty quantification with clinical data

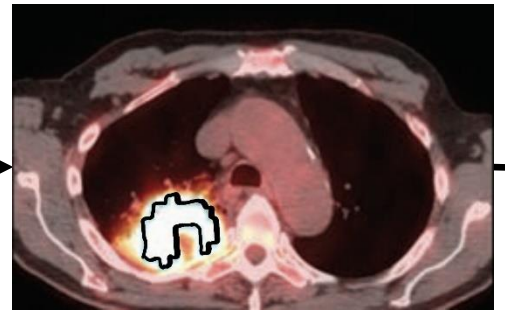


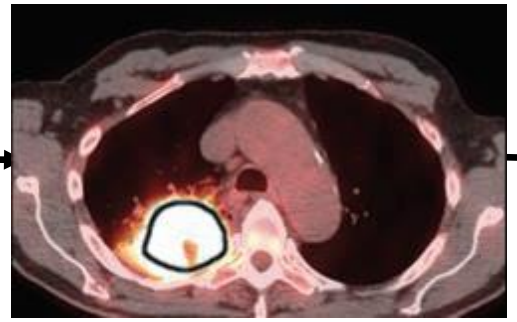
Task: Find metabolic tumor volume (MTV)

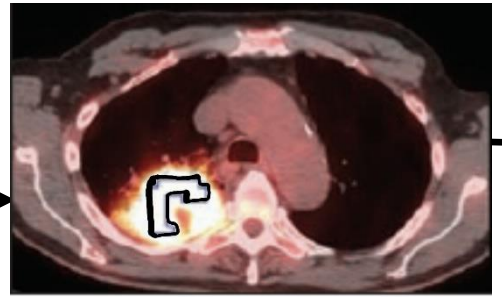
Segmentation method 1

Segm. Meth. 2

Segmentation method 3







V_1

V_2

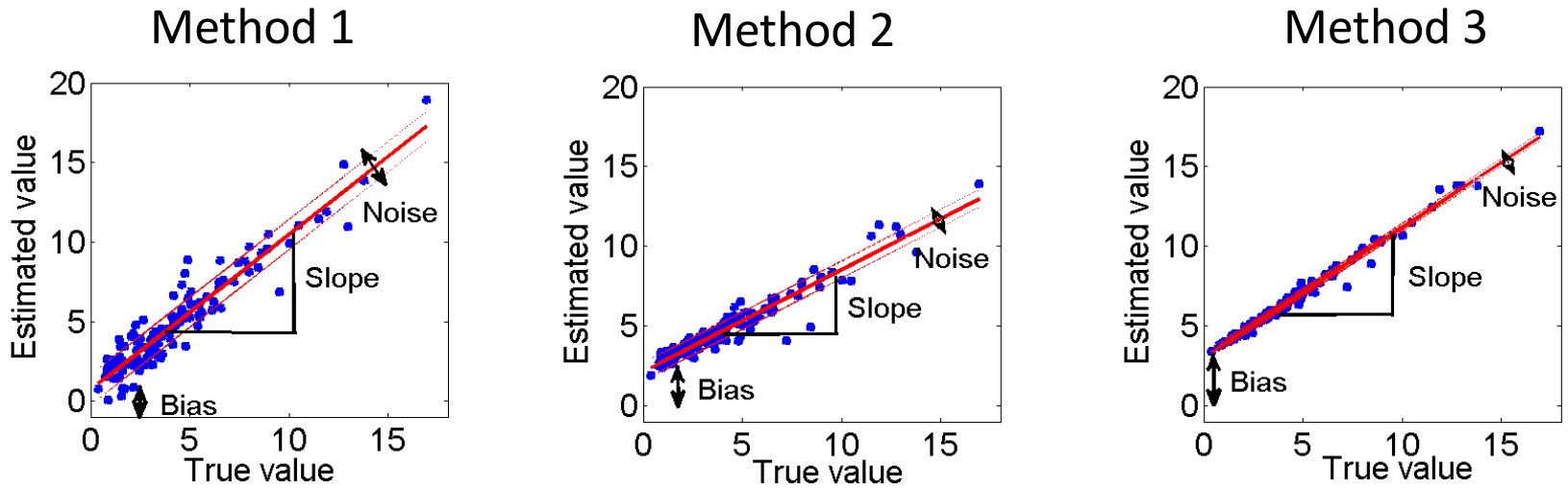
V_3

Gold standard

$V_{gold\ standard}$

Which method yields most certain values when no ground truth is unavailable?

No-gold-standard evaluation (NGSE): Premise



Assumption 1: True and estimated quantitative values related linearly by a slope, bias and noise standard-deviation term

Assumption 2: True values sampled from a parametric distribution function

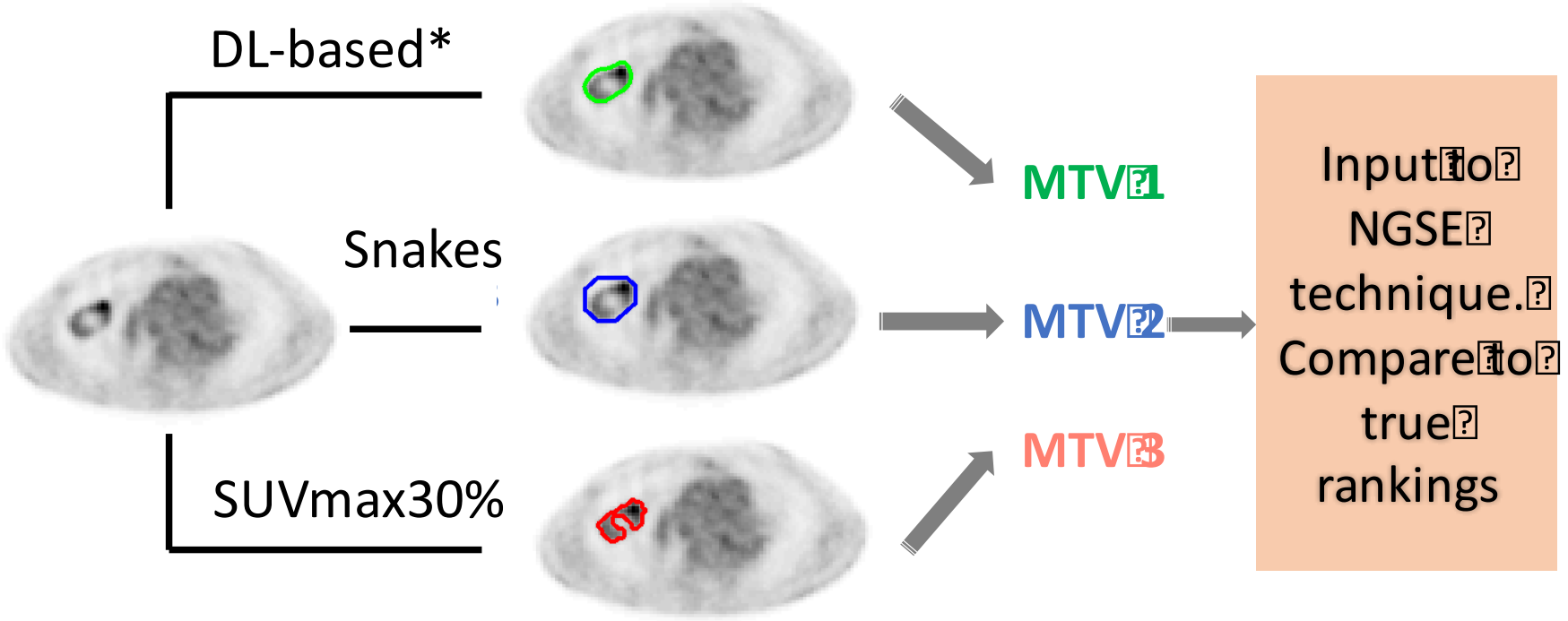
No-gold-standard evaluation (NGSE) idea: Estimate the slope, bias, and noise standard-deviation terms without knowledge of the true quantitative values[1-4]

[1]Hoppin et al 2001, Trans. Med. Imag., [2]Jha et al 2016, Phys. Med. Biol.
[3]Jha et al, J. Med. Imag. 2017 [4]Liu et al, Proc. SPIE Med. Imag. 2022

Example study on validating the no-gold-standard evaluation (NGSE) technique

Ziping Liu, PhD

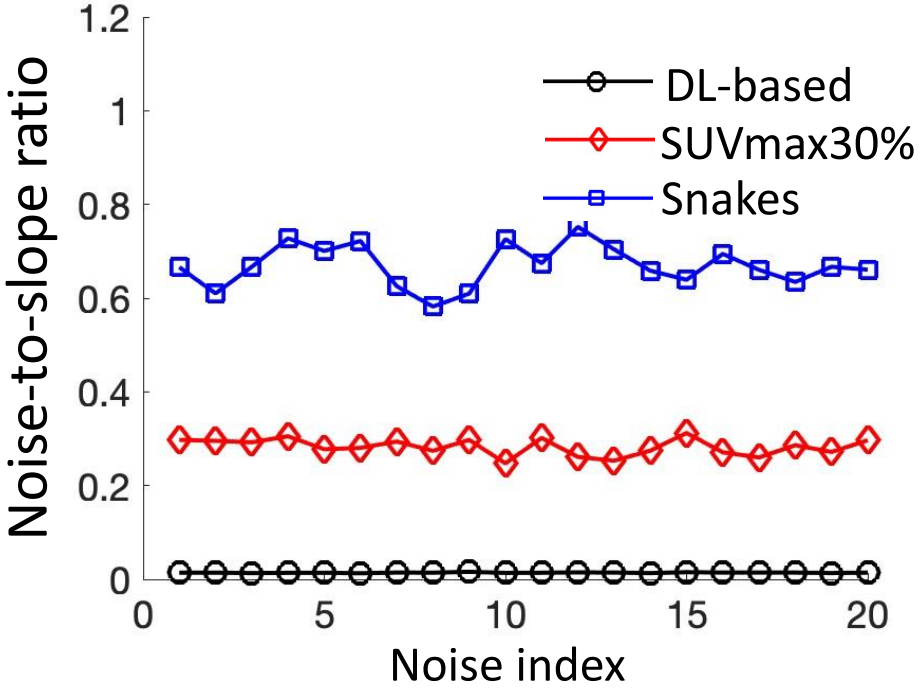
Evaluate PET segmentation methods on the task of measuring metabolic tumor volume (MTV)



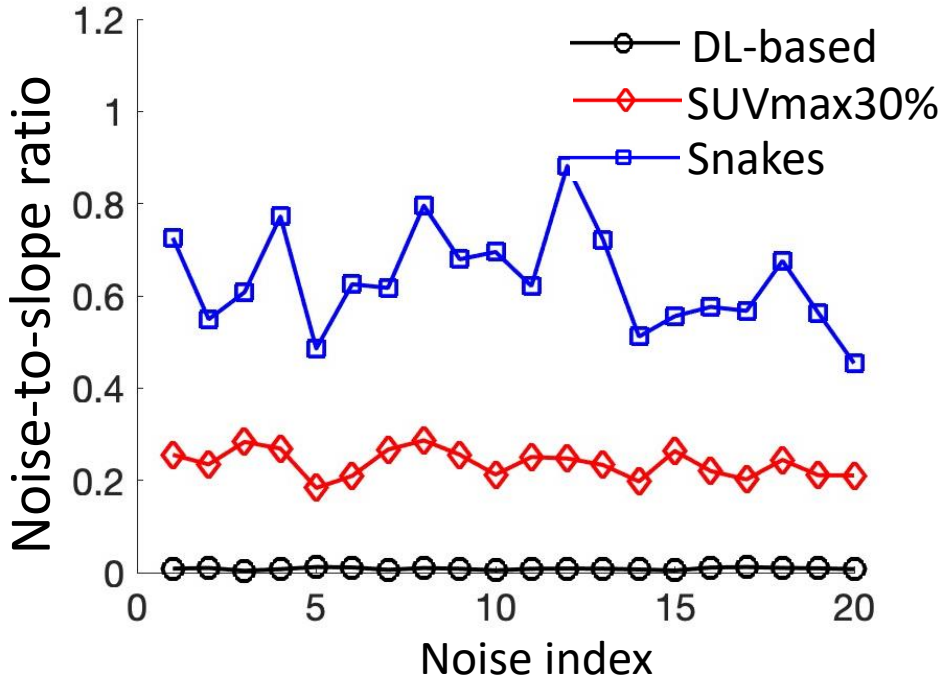
*Liu et al, Phys. Med. Biol., 2021

Results: Validation of the NGSE technique

Ground truth known



NGSE technique
Ground truth unknown



Proposed NGSE technique yielded the same rankings of the segmentation methods as when ground truth was available*

We have demonstrated similar efficacy for multiple other medical imaging modalities and developed methods to guide use of NGSE techniques**

*Liu et al, J. Nuc. Med. (suppl)

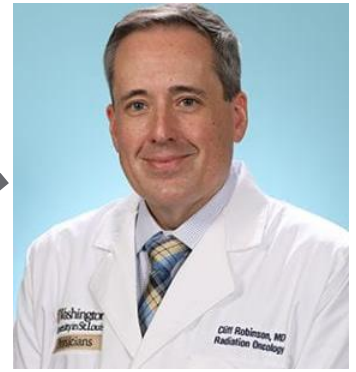
*Liu et al, Proc. SPIE Med. Imag. 2024

Aim 2: Designing survey to elicit patient risk-value profiles

Anya Plutynski, PhD
Biomedical Ethicist

Tyler Fraum, MD
Radiologist

Designing a survey to elicit patient risk-value profies in using AI for analysis of their scans



Clifford Robinson, MD
Radiation oncologist

Abhinav Jha, PhD
Computational imaging scientist

Survey design

- **Context**: Using AI-based tools for delineation of tumors from positron emission tomography/computed tomography (PET/CT) scans in patients with stage III non-small cell lung cancer
- **Goal**: Assess the patient's risk-value profiles about the use of AI-based tools in analysis of their PET/CT exams, with the objective that this could help physicians in modeling these values
- **Structure**: Educational introduction followed by questions to gauge patient attitudes about AI, risk-value profiles, and uncertainty
- **Next steps**: Finalize survey and circulate amongst patient advocates to obtain feedback and refine survey questions
- **Outcome**: A survey to elicit patient risk-value profiles to incorporate AI uncertainty

Efforts at a community level

As member of the Society of Nuclear Medicine and Molecular Imaging task force, we published papers on ethical considerations for AI in medical imaging

Scholarly work

1. J. Birch, K. A. Creel, **A. K. Jha**, A Plutynski, “Clinical decisions using AI must consider patient values”, Nature Medicine, 28, pp. 229–232, Jan 2022, PMID: 35102337 (link)
2. Y. Liu and A. K. Jha, “How accurately can quantitative imaging methods be ranked without ground truth: an upper bound on no-gold-standard evaluation”, SPIE Med. Imag. 2024 (Finalist: Robert F. Wagner Best Student Paper award)
3. J. Herington, M. D. McCradden, K. Creel, R. Boellaard, E. C. Jones, P. KC, **A. K Jha**, A. Rahmim, P. J.H. Scott, J. Sunderland, R. L. Wahl, S. Zuehlsdorff, and B. Saboury, “Ethical Considerations for Artificial Intelligence in Medical Imaging: Data Collection, Development, and Evaluation”, J. Nuc. Med., 64(12), 2023, PMID: 37827839 (link)
4. J. Herington, M. D. McCradden, K. Creel, R. Boellaard, E. C. Jones, P. KC, **A. K Jha**, A. Rahmim, P. J.H. Scott, J. Sunderland, R. L. Wahl, S. Zuehlsdorff, and B. Saboury, “Ethical Considerations for Artificial Intelligence in Medical Imaging: Deployment and Governance”, J. Nuc. Med., 2023, 64 (10) 1509-1515 (link)
5. Z. Liu, Z. Li, J. C. Mhlanga, B. A. Siegel, **A. K. Jha***, “No-gold-standard evaluation of quantitative imaging methods in the presence of correlated noise”, Proc. SPIE Med. Imag. 2022, PMID: 36465994 (link) (Finalist: Robert F. Wagner Best Student Paper award)
6. ”Measuring, defining, and reframing uncertainty in AI for clinical medicine”, final stages of preparation

Acknowledgements





Many thanks to support from NIH R01 EB031051-S2.
Also thanks to NIH R01 EB031051, R01 EB031962, NSF CAREER, NIBIB Trailblazer award
This has been a fantastic research experience!