Nuclear Medicine from a Novel Perspective: Buvat and Weber Talk with OpenAI's ChatGPT

Irène Buvat, PhD[1], and Wolfgang Weber, MD, PhD[2]

[1]Inserm Laboratory of Translational Imaging in Oncology, Institut Curie, Orsay, France; [2]University Hospital, Technische Universität München, Munich, Germany

Irène Buvat, PhD, Centre National de la Recherche Scientifique director of research and head of the Inserm Laboratory of Translational Imaging in Oncology at the Institut Curie, Orsay (France), and Wolfgang Weber, MD, PhD, director of the Department of Nuclear Medicine at "Klinikum rechts der Isar" (University Hospital of the Technische Universität München, Germany) talked with ChatGPT, an artificial intelligence (AI) language model. ChatGPT was developed by OpenAI (San Francisco, CA), a research organization dedicated to advancing artificial intelligence (AI) in a safe and beneficial way. The bot was trained on a large dataset of text, including books, articles, and other sources of information, to develop language processing and understanding capabilities. It has been designed to be able to respond to a wide range of questions and provides information quickly and easily on a variety of topics. ChatGPT was launched on November 30, 2022, and, by January 2023, had reached more than 100 million users—a faster rate of adoption than Instagram or TikTok. By February, ChatGPT had been featured in 10 articles in the *New York Times* and been on the covers of *TIME* magazine and the *Economist*. Drs. Buvat and Weber asked ChatGPT about the current practice and future development of nuclear medicine.

Commentary, Drs. Buvat and Weber: ChatGPT no longer needs an introduction for most people, but, for those who have been hibernating for the last few months, it is a "bot," a program that can interact with human users. The user types in a question in a natural language (with >90 available languages), and ChatGPT responds in natural language. The bot can be asked to adopt a specific style, such as Shakespearean or journalistic language. In conversations with users, ChatGPT always "pays attention" and remembers the thread of the dialogue. It is sometimes a bit stubborn but very polite and eager to help. Its knowledge is impressively broad, including aspects of medicine and medical imaging. When it cannot find an answer, it says so. Sound like the perfect personal assistant? Not quite. It is more like a trainee with impressive skills and knowledge. In January 2023 we challenged ChatGPT with a few questions relevant to nuclear medicine to illustrate the role it could play in our field. (**Note:** ChatGPT received no special preparation for this interview and responded entirely with information already in its extensive dataset. Bot responses to questions have not been edited to this journal's style; instead, they are reproduced exactly as they appeared to the interviewers.)

Q1, Dr. Weber: Hello. Could you please write a PET/CT report for a patient with suspected lung cancer? The scan shows a hypermetabolic mass in the right upper lobe and hypermetabolic lymph nodes in the right hilar and left lower paratracheal region. Please include the TNM stage.

ChatGPT:

Patient: [Patient Name] Date of Exam: [Date].

Indication: Evaluation of suspected lung cancer.

Technique: Whole body PET/CT scan was performed with the administration of [dose] of FDG.

Findings:

1. Hypermetabolic mass is identified in the right upper lobe (SUV max 8.2) with an associated spiculated margin, consistent with a primary lung malignancy.
2. Hypermetabolic lymph nodes are identified in the right hilar (SUV max 6.5) and left lower paratracheal (SUV max 6.7) regions, consistent with nodal involvement by lung cancer.
3. CT scan of the chest also confirms the above-mentioned findings and also shows no evidence of distant metastasis.

TNM stage: T3N2M0 (The tumor is located in the right upper lobe, it is larger than 7cm and there is nodal involvement in the right hilar and left lower paratracheal region, no evidence of distant metastasis).

Impression: Hypermetabolic mass in the right upper lobe with associated spiculated margin and hypermetabolic lymph nodes in the right hilar and left lower paratracheal region consistent with lung cancer. Recommend further workup and histopathological correlation, including biopsy of the primary tumor and involved lymph nodes.

Commentary, Drs. Buvat and Weber: This quite impressive response was given within a few seconds! ChatGPT uses the typical language of radiology reports and provides SUV$_{max}$ values that

Irène Buvat, PhD

Wolfgang Weber, MD, PhD

are consistent with high metabolic activity. It also gets the TNM stage almost right. It interprets the left paratracheal nodes as "contralateral mediastinal nodes" and assigns the correct N2 stage. Interestingly, it uses the slightly outdated 7th edition of lung cancer TNM staging, which defines T3 as a tumor with a diameter of >7 cm. The current 8th edition defines T3 by a tumor diameter between 5 and 7 cm.

Commentary, Dr. Weber: Although ChatGPT should definitely not be used to write patient reports at this stage, one could envision that an AI could be trained specifically for writing radiology reports in the near future. Such a system could automatically generate structured reports and avoid all the common typographic and grammatic errors that plague current radiology reports. It also does not seem to be too far-fetched to combine an AI that analyzes images with AI that generates a human-readable, standardized report. Such a system could dramatically streamline radiology workflows if appropriately validated and would enable radiologists to report many more scans than currently possible.

Q2, Dr. Buvat: Could you please explain what the Ki Patlak slope calculated from an FDG PET/CT scan represents?

ChatGPT: The Ki Patlak slope from FDG PET/CT scan represents the rate of influx of the radiotracer (fluorodeoxyglucose, FDG) into tissues, which reflects the local tissue glucose metabolism. It is used to quantify the regional cerebral blood flow (rCBF) and can provide information about tissue perfusion and functional brain activity. The Ki Patlak method is a graphical analysis that plots the distribution volume of the tracer in tissue over time. The slope of this line

most accurate and specific measure of metabolic activity within a lesion.

Commentary, Dr. Buvat: Here we see the limits of ChatGPT's knowledge. The sentences sound correct, but the content is mostly wrong. This illustrates one of the dangers of ChatGPT: its responses can sound very convincing even when they are simply wrong.

Q4, Dr. Buvat: Based on current knowledge, what do you think are the 3 most promising radiotracers in oncology for routine use in the future?

ChatGPT: There is ongoing research and development in the field of molecular imaging and radiotracers, and several promising radiotracers are being studied for their potential use in oncology. Here are three radiotracers that are considered to be among the most promising for routine use in the future:

PSMA-PET: PSMA is a protein that is highly expressed in prostate cancer, and PSMA-PET imaging using radiolabeled PSMA ligands has shown high sensitivity and specificity for detecting prostate cancer, even in its early stages.

68Ga-DOTA-peptides: 68Ga-DOTA-peptides are a group of radiotracers that bind to specific receptors or antigens on cancer cells and are used for imaging and staging of various types of cancer, including neuroendocrine tumors, gastrointestinal tumors, and others.

18F-FDG: 18F-FDG is a widely used radiotracer that has been validated for the evaluation of malignant lesions, including various types of cancer such as lymphoma, lung cancer, and others. The high glucose uptake in malignant cells makes 18F-FDG a useful tool for detecting and staging cancer.

"[O]ne of the dangers of ChatGPT: its responses can sound very convincing even when they are simply wrong."

---

represents the rate of tracer uptake, with a steeper slope indicating a faster uptake rate and thus higher tissue metabolism. This method is commonly used in the assessment of neurological and oncological disorders, among other applications.

Commentary, Dr. Buvat: An almost correct answer. ChatGPT mixes up cerebral blood flow and cerebral metabolic rate in 1 sentence. The information is probably more comprehensive than a Google search about Patlak analysis. It is remarkable that ChatGPT can answer this question, because Patlak analysis is used only in nuclear medicine research, which constitutes only a very small part of the overall information found on the Internet.

Q3, Dr. Buvat: In a tumor seen on an FDG PET/CT image, would you recommend reporting SUV_{peak} or SUV_{max} in the medical report?

ChatGPT: In a medical report for an FDG PET/CT scan, both the SUVpeak and the SUVmax can be reported as measures of tumor metabolic activity. However, the SUVpeak is typically used as the standard measure in clinical practice.

SUVpeak refers to the maximum standard uptake value (SUV) within a region of interest (ROI) that is used to define a specific lesion, whereas the SUVmax refers to the highest SUV within the entire image. The SUVpeak provides a more accurate representation of the metabolic activity within a lesion, as it takes into account only the highest SUV within that lesion, rather than including the background or normal physiological activity.

Therefore, it is generally recommended to report the SUVpeak in the medical report for an FDG PET/CT scan, as it provides the

It is important to note that the use of these and other radiotracers is dependent on ongoing research, regulatory approval, and clinical validation, and the most promising radiotracers for routine use may change over time as new information becomes available.

Commentary, Dr. Weber: A nicely worded answer to what has literally become a "billion-dollar question." ChatGPT cannot see into the future but summarizes nicely what has been said about the 3 most commonly used classes of tracers. It doesn't really answer the question, because all these agents are already in clinical use. But, of course, their use may increase in the coming years, and, therefore, the answer is not incorrect. A human "expert" would probably have mentioned agents that are not yet FDA approved and are undergoing clinical testing, e.g. fibroblast-activation protein ligands or gastrin-releasing peptide-receptor ligands.

Q5, Dr. Buvat: Do you think I can trust AI-based segmentation tools for segmenting head and neck tumors from PET/CT in the context of radiation therapy planning?

ChatGPT: AI-based segmentation tools for head and neck tumor segmentation from PET/CT scans have shown promise and have the potential to improve the accuracy and efficiency of radiation therapy planning. However, their reliability and accuracy depend on the quality of the training data used to develop the AI algorithms, as well as the complexity and heterogeneity of the tumors being segmented.

Currently, AI-based segmentation tools are often used as a tool to support human interpretation and decision-making, rather than as a replacement for human expertise. It is important to validate the results

of AI-based segmentation tools against established clinical criteria and to assess their performance in a robust and systematic manner.

In conclusion, AI-based segmentation tools have the potential to be a valuable tool in radiation therapy planning, but they should be used in conjunction with human expertise and with appropriate validation to ensure the reliability and accuracy of the results.

Commentary, Dr. Buvat: This answer is again quite impressive; there is little a human expert could add. The key challenges are described (i.e., heterogeneity of the tissue, need for sufficient training sets, etc.).

Q6, Dr. Buvat: Which such tool would you recommend?

ChatGPT: As a language model developed by OpenAI, I don't have personal opinions or biases, and I do not endorse any specific product or brand. The choice of an AI-based segmentation tool for head and neck tumor segmentation from PET/CT scans depends on several factors, including the clinical and technical requirements of the specific case, the available resources, and the preferences and experience of the clinician.

When selecting an AI-based segmentation tool, it is important to consider the following factors:

- Validation and reliability: The tool should have been validated against established clinical criteria and should have a demonstrated track record of reliability and accuracy.
- User-friendliness: The tool should be easy to use and should provide results in a timely and efficient manner.
- Technical support and customer service: The vendor should provide robust technical support and customer service to assist with any questions or problems that may arise during use.
- Integration with other software and systems: The tool should be compatible with the other software and systems used in the clinical environment, such as the radiation therapy planning system, and should provide seamless integration.

Ultimately, the choice of an AI-based segmentation tool should be based on a thorough evaluation of the specific requirements and resources of the clinical environment, and should be made in consultation with relevant experts and stakeholders.

Commentary, Dr. Weber: A very useful answer that avoids giving a concrete recommendation. All the factors ChatGPT lists are relevant when deciding which computer system to purchase. It would be difficult for a human expert to list all these items as quickly as ChatGPT did. This illustrates how ChatGPT can be helpful for collecting relevant information before making a purchase decision.

Q7, Dr. Buvat: Do you think it is the right time to build a theranostic center for molecular radiation therapy?

ChatGPT: The development of a theranostic center for molecular radiation therapy is a complex and multi-disciplinary effort that requires significant resources and expertise. Whether it is the "right time" to build such a center depends on several factors, including the local demand for advanced cancer treatment options, the availability of expertise and resources, and the regulatory and reimbursement landscape.

In general, the field of molecular radiation therapy is rapidly evolving and has shown significant promise in improving the outcomes of cancer treatment. There is increasing demand for these advanced treatment options, particularly among patients with complex or advanced-stage cancers.

However, the development of a theranostic center is a significant investment and requires a commitment to long-term sustainability, as well as a comprehensive understanding of the regulatory, reimbursement, and economic factors that impact the delivery of these treatments.

In conclusion, the "right time" to build a theranostic center for molecular radiation therapy will depend on the specific circumstances and resources of the local environment. It is important to carefully assess the feasibility and sustainability of such a project, and to consult with relevant experts and stakeholders before proceeding.

Commentary, Dr. Weber: ChatGPT does very well with this timely question. It describes the potential of a theranostic center but also mentions the important regulatory and economic hurdles that physicians will face when setting up such centers. A human expert might have added that the current interest in theranostics almost exclusively stems from 2 diseases: neuroendocrine tumors (NETs) and prostate cancer. Alternative therapies for prostate cancer or NETs that are more effective, have less side effects, or are less expensive could significantly lessen the clinical relevance of theranostics. It is also largely unclear how quickly theranostic targets for other diseases can be identified that would broaden the spectrum of a theranostic center and decrease its dependence on only 2 therapeutic targets.

Summary commentary, Drs. Buvat and Weber: Overall, these chats illustrate the strengths and weaknesses of current bots. They can certainly help in time-consuming tasks, such as writing medical reports (Q1) or creating cover letters when submitting articles or simplified summaries for the public based on scientific abstracts. They can answer specific technical questions well and be used for educational purposes (Q2), but answers should always be verified, because they can look authoritative but be wrong (Q3: the definition of SUV_{peak} is approximate and the definition of SUV_{max} is not quite correct). ChatGPT has learned from data collected through 2021 and is clearly not aware of the latest advances in our field (Q4, fibroblast-activation protein is missing from the list) and does not fully understand the question (Q4: FDG is not a promising radiotracer). The bot is very cautious about giving recommendations (Q5) and does not promote a solution (Q6), as a colleague might do. It offers a synthetic view on current issues in only a few seconds (Q7) and avoids taking a firm position (Q3, Q7). Thanks to fierce competition among the GAFAM (Google, Apple, Facebook, Amazon, Microsoft), chatbots will evolve very quickly and will become more and more reliable as personal assistants. Thanks to continuous learning techniques, they will expand their knowledge and improve the accuracy of their answers. They should not be feared but rather considered promising allies who will help us cope with ever-increasing workloads, freeing up time to devote to patients and colleagues and to expend more energy on tasks that require unique and advanced expertise. A final note: ChatGPT is often quite busy and not always available. Sound familiar?