Published in final edited form as:

PET Clin. 2022 January ; 17(1): 145–174. doi:10.1016/j.cpet.2021.09.006.

Artificial Intelligence in Lymphoma PET Imaging:: A Scoping Review (Current Trends and Future Directions)

Navid Hasani[1], **Sriram S Paravastu**[2], **Faraz Farhadi**[2], **Fereshteh Yousefirizi**[3], **Michael A Morris**[4], **Arman Rahmim**[5], **Mark Roschewski**[6], **Ronald M Summers**[7], **Babak Saboury**[8]

[1]Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Building 10, Room 1C455, Bethesda, MD 20892, USA; University of Queensland Faculty of Medicine, Ochsner Clinical School, New Orleans, LA 70121, USA.

[2]Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Building 10, Room 1C455, Bethesda, MD 20892, USA.

[3]Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC, Canada.

[4]Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Building 10, Room 1C455, Bethesda, MD 20892, USA; Department of Computer Science and Electrical Engineering, University of Maryland-Baltimore Country, Baltimore, MD, USA.

[5]Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC, Canada; Department of Radiology, BC Cancer Research Institute, University of British Columbia, 675 West 10th Avenue, Vancouver, British Columbia, V5Z 1L3, Canada.

[6]Lymphoid Malignancies Branch, Center for Cancer Research, National Institutes of Health, Bethesda, MD, USA.

[7]Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Building 10, Room 1C455, Bethesda, MD 20892, USA.

[8]Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Building 10, Room 1C455, Bethesda, MD 20892, USA; Department of Computer Science and Electrical Engineering, University of Maryland-Baltimore Country, Baltimore, MD, USA; Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Abstract

Malignant lymphomas are a family of heterogenous disorders caused by clonal proliferation of lymphocytes. ^{18}F-FDG-PET has proven to provide essential information for accurate quantification of disease burden, treatment response evaluation, and prognostication. However, manual delineation of hypermetabolic lesions is often a time-consuming and impractical task. Applications of artificial intelligence (AI) may provide solutions to overcome this challenge. Beyond segmentation and detection of lesions, AI could enhance tumor characterization and heterogeneity quantification, as well as treatment response prediction and recurrence risk stratification. In this scoping review, we have systematically mapped and discussed the current applications of AI (such as detection, classification, segmentation as well as the prediction and prognostication) in lymphoma PET.

Keywords

Artificial intelligence; Deep learning; Detection; Lymphoma; Positron emission tomography (PET); Radiomics; Radiophenomics; Segmentation

Introduction

Lymphomas are a diverse group of hematologic malignancies, which can be broadly categorized into Hodgkin (HL) and non-Hodgkin diseases (NHL) and have a wide range of clinical presentations.[1] 2-deoxy-2-[Fluorine-18]fluoro-D-glucose (^{18}F-FDG) positron emission tomography/computed tomography (PET/CT) is extensively used for staging and response assessment in HL and NHL.[2, 3, 4, 5] The accurate and precise quantification of tumor burden in lymphoma is critical for prognosis and treatment response evaluation and prediction. ^{18}F-FDG PET scans provide valuable information about the metabolism of lesions. This functional information combined with structural (CT or MRI) data can be used to assess the global disease burden in Alzheimer's disease,[6] Crohn's disease,[7] knee inflammation[8] as well as lymphoma.[9] Therefore, ^{18}F-FDG PET/CT is extremely valuable in the noninvasive assessment of disease burden.[10]

To determine global disease burden, segmentation of all tumor lesions is a vital step that allows the measurement of metabolically active tumor volumes (*MTV*), mean activity of the lesion (SUV_{mean}), lesion partial volume corrected metabolic volume product (*PVC-MVP*: calculated as the product of *lesion MTV* and *lesion PVC-*SUV_{mean}), total metabolic tumor volume (*TMTV*: calculated as the sum of MTV of all lesions), total lesion glycolysis (*TLG*),[11] whole-body metabolic burden (*WBMB*: calculated as the sum of *lesion PVC-MVP* of all lesions),[10,12] metabolic heterogeneity (*MH*)[13,14] and lesion dissemination (D_{max}).[15,16] There are, however, *various methods for* the *segmentation* of tumor lesions (e.g. manual, thresholding-based, region-based, or boundary-based)[17,18] each with high *inter-observer variability* depending on the operator and segmentation method.[19,20] Furthermore, even for an expert, manual segmentation takes time (30–45 minutes per patient depending on tumor burden) as summary measurements of each lesion must be aggregated.[19] Lymphoma lesion segmentation is a challenging task due to the large variability in number, size, distribution, uptake, the shape of lesions, and different degrees of glucose metabolism (Fig. 1)[21, 22, 23] Normal biodistribution of ^{18}F-FDG creates physiologic intense activity either due to high metabolic rate (such as brain) or high concentration of excreted radiotracer (such as renal collecting ducts and bladder). This normal pattern significantly deteriorates the performance of the crude intensity-based detection and segmentation methods.[24,25] Automated segmentation and feature extraction approaches may be an exciting avenue to limit measurement discrepancies and cut image analysis to a fraction of the current required time.[19]

In general, the most critical aspects that should be evaluated during lymphoma PET are as follows: (1) quantification of disease burden,[26,27] (2) evaluation of therapy response,[28] and (3) extraction of additional image information used for prognosis and diagnosis of lymphoma.[29,30] Artificial Intelligence (AI) has ample potential to achieve

the aforementioned goals by first performing automatic quantification, which entails (1) automatically identifying the location of the abnormality,[31] (2) automatically segmenting the lesion,[32,33] (3) summarizing each lesion to other dimensions (SUV_{mean}, SUV_{max}).[10,34] Finally, AI can enable registration at multiple points in time,[35,36] scaling from one space to another. This allows evaluation of lymphoma before and after diagnosis or therapy.[37,38] Given such significant promise that AI has in other fields of medical imaging and sporadic relevant evidence specific to lymphoma PET, it is important to scope and map current applications of AI in lymphoma PET imaging. However, at the time of this publication, there has not been a review of various applications of AI as it pertains to lymphoma.

Thus, in this article, we first aim to scope the breadth of evidence and systematically map literature on the topic of AI applications in lymphoma PET to identify key concepts and disseminate research evidence on various AI models. We then depict the potential clinical utility of AI in PET imaging and anticipate future directions that can be expected for AI applications in lymphoma.

A list of abbreviations used in this article are shown in Box 1.

Methods

This scoping review was conducted following the preferred reporting items for systematic reviews and meta-analysis extension for scoping reviews (PRISMA-ScR) guidelines.[39]

Search Strategies

Bibliographic searches were performed in PubMed, EMBASE, Cochrane Library, and Google Scholar for articles published before September 1st, 2021. In PubMed/Medline, Medical Subject Headings (MESH) in all fields were searched for “Artificial Intelligence” (or Deep Learning or Machine Learning or Support Vector Machine (SVM) or Convolutional Neural Network (CNN) or Artificial Neural Network (ANN)) and “Positron Emission Tomography” (PET or PET-CT or PET-MR) and “lymphoma.” The remainder of the studies were identified through manual searches of bibliographies and citations until no further relevant studies were found. One investigator (N.H) independently screened titles and abstracts and selected relevant citations for full-text review.

Eligibility and Exclusion Criteria

Studies that reported the diagnostic measurement of an AI/ML/DL algorithm to investigate any type of lymphoma using PET were sought. Articles were excluded if the study was not written in English. All the nonpeer-reviewed material such as nonpeer-reviewed conference articles and archives as well as studies irrelevant to applications of AI in lymphoma PET imaging were excluded. Studies that comprised a development or assessment of an AI algorithm on PET imaging in human populations diagnosed with lymphoma or any subtypes were eligible. The search included all primary articles since the beginning of 2009.

Data Extraction and Analysis

Key study characteristics (such as tasks/models, methods, results) for selected papers are summarized. According to the specification of the task—segmentation, classification, prognostication—the articles were categorized (Table 1). The details of methods and the AI architecture proposed were recorded. Study characteristics extracted were the purpose of the article, authors, year of publication, AI model design, proposed AI application, and ground truth (GT). Also extracted were information regarding sample size, training sample, testing, and validation samples as well as figures of merit (FoM) such as specificity, sensitivity, dice similarity coefficient (DSC), and Hausdorff distance (HD) (Table 2) depending on the proposed application of the algorithm (see Table 1).

Results

Search Results

We retrieved 1122 documents from initial searches; 1089 met the eligibility criteria for the title and abstract review and 75 met the final criteria for full-text review as shown in Fig. 2. After the screening process, 20 articles were included; these covered both the AI development and clinical assessment fields (see Fig. 2). All 20 papers examined AI applications in lymphoma PET imaging. These studies either developed an original model or evaluated a previously proposed AI model to perform detection, classification, segmentation, characterization, prediction/prognosis, or a combination of these tasks on PET/CT or PET images. The definition of the aforementioned tasks is provided in “Terminology for Elucidating Algorithm Aim” under the Results section.

Overview: Key Literature Characteristics

In this section, we will systematically examine how each study performed these tasks assigned to AI. We first identify the reported AI task (detection, segmentation, classification, radiophenomics), then determine the model’s input and output for that specific task. For instance, regarding the detection task, we identify the AI’s input, which is frequently in the form of pixels, and the output would be detecting areas suspicious for cancer (high FDG uptake). Table 1 summarized the results of each study as it pertains to the proposed model (such as CNN, ANN, and so forth), task (classification, detection, segmentation, prediction, prognosis, and so forth), FoM provided specific to each task (see Table 2), and the GT definition for evaluating the results. Based on the current literature, the 2 main applications for radiomics in lymphoma: distinguishing lymphomas as separate form other tumors, and prediction or prognostication of lymphomas.[40]

Appropriate GT and label for an imaging AI application are highly related to AI objectivity. PET images are often visually analyzed, and this may often lead to high inter-and intraoperator variability. Thus, it is a challenge to define optimal GT for datasets to be used for AI training, and a suboptimal GT will hamper the predictive accuracy of the model.[41,42] For these reasons, here, we present a definition of GT specific to the AI objective as provided by each of the studies (see Table 1).

Terminology for Elucidating Algorithm Aim

Detection—Detection as a task refers to locating an area within an image that contains an object of interest with a stated level of certainty. This task often involves a combination of localization and some level of classification, finding a nodule in the lung is an example of a detection task. As referred to in Table 1, the input to an AI algorithm that performs detection should be a type of image (pixel/voxel), and the output should also be a location containing the object of desire.[24] For example, Bi and colleagues use 3D WB, coronal ^{18}F-FDG PET with CT slices (2 channels) as an input to detect individual regions of High Normal Activity (HiNA) (also referred to as sites of FDG excretion and physiologic FDG uptake (sFEPU) by Bi and colleagues) using their multi-scale superpixel encoding CNN model.[24] In addition, Sibille and colleagues proposed a model that automatically detects HiNA and lesions suspicious of lung cancer and lymphoma lesions[50] (Fig. 3). Similarly, Wiseman and colleagues proposed a 3D DeepMedic model that implicitly learned information about the HiNA regions during model training achieving 85% detection TPR on average.[49] Due to the increased heterogeneity of HiNA regions below the diaphragm (for example in bladder, kidney, and ureter) this model performed better above the diaphragm than the below.[49]

In this scoping review, we found 7 studies for lymphoma lesion detection in PET/CT imaging (see Table 1 - Yuan and colleagues (2021),[46] Zhou and colleagues (2021)[38] Weisman and colleagues (2021),[61] Sibille and colleagues (2020),[50] Hu and colleagues (2019),[56] Yu and colleagues (2018),[57] Bi and colleagues (2017)[24]).

Segmentation—Delineation of the boundary of an *object of interest* given its location is referred to as segmentation.[62] Accurate segmentation of lymphoma is an important task as it permits the extraction of both lesion-level (such as SUV_{max} and SUV_{mean}) and whole-body quantitative metrics (such as TMTV) which provide important predictive and prognostic information.[63,64] The image input data for the segmentation task can be as large as a 3D WB PET image or as small as a group of pixels.

The segmentation task which is the combination of localization and pixel/voxel level classification often has two inputs (1) the image that contains an *object of interest* (input 1) (2) the location of that *object of interest* with a certain level of certainty (input 2). Input 2 can either be a probability map of coarse region of the lesions, or single-pixel locations representative of the lesion locations which can be derived from a localization or a detection task. The *output* of segmentation is an image that encodes the membership of each voxel to the *object of interest* (this output can be referred to as a segmentation map).

The second input for the tumor segmentor can be entered manually (eg, Sadik (2019)[52]) or automatically (eg, Yuan and colleagues)[46] into the system. In case of automatic input, the detection probability map results can be fed into the algorithm. This can be called a *cascaded approach* whereby all of the steps are performed separately and sequentially by one or several neural networks (Fig. 4). These techniques often divide the segmentation process into detection, and segmentation phases and provide specific evaluation metrics for each step along the way.

In this scoping review, six studies performed the segmentation task (see Table 1- Pinochet and colleagues (2021)[43] Yuan and colleagues (2021)[46] Sadik (2019)[52] Hu and colleagues (2019)[56] Yu and colleagues (2018)[57] Grossiord and colleagues (2017).[58])

End-to-end segmentation—When the segmentation task is done in one step, we use the term end-to-end segmentation in which case the only input is the image data. These approaches optimize for efficiency and performance in terms of memory consumption and in case of limited access to well-annotated training data.[65] An example, Weisman and colleagues proposed an end-to-end segmentation model that receives WB ^{18}F-FDG PET and CT image through 2 channels and without an additional detection step, the CNN is able to provide a map of masked segmented lymphoma lesions with DSC of 0.86 (Fig. 5).[48]

In this review, there were 4 studies that performed the end-to-end lesion segmentation task (see Table 1 - Blanc-Durand and colleagues (2020),[47] Weisman and colleagues (2020),[48] Li and colleagues (2019),[51] and Desbordes and colleagues (2016)[59]).

Classification—Through reviewing the literature, we recognized 2 distinct usages of the word classification. To avoid ambiguity, we have to clarify these terms here: in a mathematical and statistical context, categorization of members of a set to various classes is defined as “classification.” This is a broad meaning of this term, and we refer to this as “statistical classification.” However, in computer vision, classification has a narrower meaning. It refers to the categorization of an image. We refer to this meaning as “image classification.” For example, we refer to the process of using radiomic feature inputs to assign patients to different diagnostic or prognostic groups by the term “statistical classification” (Refer to Section “Prediction and Prognosis” under Results section). For example, Annunziata and colleagues[22] used image features such as Deauville Score, qPET, MTV_0 from end-of-treatment ^{18}F-FDG PET, and CT to classify the prognosis of patients into “relapse” or “progression” classes. In contrast, the process of categorizing an image input (such as the axial slice of the PET) to normal versus abnormal is “image classification.” For example, Lippi and colleagues used a machine-learning algorithm that classified the lesions within the image into 4 malignant lymphoma subtypes: DLBCL, HL, follicular and mantle cell lymphomas.[54] Radiomics signatures such as PET and CT textural features have shown very good performance to classify the disease sites from physiologic uptake sites and inflammatory nonlymphomatous sites. For instance, to classify an ^{18}F-FDG avid lesion (benign vs malignant), Lartizien and colleagues developed an SVM and Random Forest (RF) model based on 12 different radiomic features extracted from PET and CT scans achieving Area Under the Curve (AUC) of 0.91[60].

In this scoping review, there were 5 studies that performed image classification (see Table 1.—Pinochet and colleagues (2021),[43] Sadik and colleagues (2021),[44] Guo and colleagues (2021),[45] Lippi and colleagues (2019),[54] Grossiord and colleagues (2017),[58] and Lartizien and colleagues (2014).[60]

Prediction and prognosis—Recurrence is common in patients with HL and NHL, emphasizing the importance of risk stratification, prognostication, and relapse prediction based on PET studies. In the context of this paper, the task of prediction and prognosis

of lymphoma based on PET imaging is performed by *statistical classification*. The inputs in statistical classification are radiomic features (eg, SUV, TMTV, TLG, entropy among others). These inputs are used to inform the output in the form of a prognostic or predictive classification.[65]

Baseline TMTV can be used for risk stratification and be a prognostic factor in a range of lymphomas (DLBCL, primary mediastinal B-cell lymphoma (PMBCL) and HL).[11,66] This is exemplified by Vercellino and colleagues, whereby the authors analyzed the prognostic capability of TMTV in patients aged 60 to 80 with DLBCL and found that a high baseline TMTV indicates poorer PFS and OS.[66]

In addition to the usage of a singular radiomic feature to predict and prognose lymphoma, some studies may combine various radiomic features to perform the same task. Mayerhoefer and colleagues used both the maximal standardized uptake value (SUV_{max}) and entropy to predict survival in mantle cell lymphoma (MCL).[22,43,55] Entropy is a measure of glucose metabolism heterogeneity within the TMTV.[55] Demonstrating the concept of combination of radiomic features to predict and prognose, Mayerhoefer and colleagues, used entropy (heterogeneity of glucose metabolism) as well as TMTV and SUV_{max} for the prediction and prognosis of PFS in mantle cell lymphoma patients with 0.72 AUC.[55]

In this scoping review, we found 3 studies that performed the statistical classification of lymphoma (see Table 1. Pinochet and colleagues, Annunziata and colleagues, Mayerhoefer and colleagues[22,43,55])

Discussion

The major objective of this study is to review recent papers in the field of artificial intelligence-based PET medical imaging in lymphoma. According to our findings, the most prevalent uses of artificial intelligence in lymphoma PET imaging are presently focused on tumor burden evaluation (detection, segmentation, and advanced quantification of lesions). In our discussion, we focus on two key themes derived from our research findings. First, we review the implications for the clinical transition of AI-based applications in lymphoma patient care. Next, we cover some critical concepts that clinicians should consider when evaluating and validating AI algorithms. In addition, we offer our thoughts on the field's future directions.

Clinical Implementation of Artificial Intelligence in the Management of Lymphoma

Currently, a prominent objective in lymphoma PET image quantification is the evaluation of disease burden by TLG and TMTV, which requires the detection and segmentation of all lesions.[67,68] In this workflow, a major bottleneck toward improved prognostic pipelines and treatment planning is the segmentation.[51,69] This is in part due to the time-consuming task of manual segmentation and a high degree of intra- and interobserver variability.[70,71]

AI approaches can help by: (1) Automated detection and segmentation (*fully-automated*) and[32,58,72](2) user detection/selection of the lesion followed by AI-based segmentation (*semi-automated*). A clear advantage of a fully automated AI model is that it can further

enhance the workflow without requiring the nuclear radiologist to identify each lesion separately. Given the extent of lymphomatous involvement, individual identification of each lesion could be time-consuming, and unlikely feasible in the routine workflow given clinical demands and traditionally available resources. A few studies carried out a fully automatic disease burden assessment of lymphoma on PET images (see Table 1- Pinochet and colleagues,[43] Bi and colleagues,[24] Li and colleagues,[51] Blanc-Durand and colleagues[47]). For example, Grossiord and colleagues used RF classifier and morphologic hierarchy to first extract PET and CT image features to classify lesions into 3 categories: organ, tumor, and nonrelevant. Then, they automatically segment lesions into the tumor category.[58] Although fully automated models can remove a layer of physician lesion identification from the workflow, semiautomated models may provide other benefits in terms of accuracy or precision. Human involvement in semiautomatic methods manifests in 2 ways: (1) Presegmentation human-based lesion detection and identification + automatic segmentation of the identified lesion (2) Automatic detection and segmentation of lesions with high false-positive rate and reliance on human agents to select the real lesions and discard the false-positives. As an example of the human selection of real lesions and deletion of false-positives, Yu and colleagues[57] manually identified lesions of true lymphoma after all possible detected lesions were automatically segmented.[57]

The biodistribution of ^{18}F-FDG creates regions with HiNA (for example in kidneys, bladder, brain, and heart) which can cause inaccurate AI-based identification and segmentation of lymphoma lesions. However, to improve the performance of a model one may attempt to exclude HiNA regions from the scene before the process of lesion detection. This may be conducted manually or automatically as a pre or postprocessing step.[51,73] By removing the HiNA regions from the training data, the performance of automated AI techniques for lesion detection and segmentation can be improved. For example, Yu and colleagues used a semiautomatic approach to identify and remove HiNA regions followed by an AI algorithm to automatically detect lymphoma lesions in patients.[57] For the purpose of improving workflow and integration into clinical practice, an end-to-end lymphoma detection AI algorithm can be trained to combine the task of HiNA region identification and detection of lymphoma regions.

We previously discussed methods of lymphoma segmentation on ^{18}F-FDG PET/CT studies. Along with simplifying the clinician's workflow, automatic segmentation of lymphoma lesions can enable end-to-end prognostication and radiomic analysis of the studies to gain valuable insight into therapy augmentation, remission planning, and recurrence prediction. There are 2 avenues for downstream prediction and prognosis of lymphoma: (1) an end-to-end prognostication/prediction task (whereby the image input is processed to output a prognostication/risk stratification category directly, with no reporting of the intermediate steps; for example Sadik and colleagues),[44] or (2) a radiomic analysis based prognostication, which is carried out as a secondary step after explicit segmentation. For example, Guo and colleagues[11] employ deep radiomics by using a neural network to derive *deep features*. *Deep features*, in contrast to *handcrafted features* of classical radiomics, are not predefined; a CNN learns features of interest depending on the task at hand and the input images. Automatically segmented lesions can also be entered into a classic radiomics pipeline to derive *handcrafted features* such as SUV, volume, or entropy. These features are predefined

by formulas and can be statistically analyzed to derive prognoses or a risk stratification schema. A range of studies showcasing classification, prognostication, and prediction based on ^{18}F-FDG-PET radiomic features is presented in Table 3.

Important Considerations for Evaluating and Validating Artificial Intelligence in Lymphoma Positron Emission Tomography

The transition of AI-based technologies to patient care requires important considerations that are both general for any AI algorithm and specific to lymphoma. In addition to the accuracy of the lesion detection and segmentation, it is very important to report the amount of time clinicians should dedicate to verify and correct model outcomes. Lumped sensitivity and specificity for lesion detection are not reflective of clinical significance as some lesions are much more important than others and "critical misses" are not tolerable. In both detection and segmentation tasks, missed lesions occur particularly for the interim scans with shrunk tumors or patients with smaller lesions. TMTV, as the only figure of merit for the assessment of detection/segmentation, may undermine the smaller missed lesions. To address this shortcoming, additional performance evaluation measures such as the D_{max} and MH should be used to define a task-specific composite FoM. Weisman and colleagues characterized the results of their end-to-end segmentation using SUVmax, MTV, TLG, SA/MTV, D_{max} to depict a well-rounded assessment of the CNN performance.[48]

As demonstrated in Table 1, the studies use a range of different methods to determine their GT based on which the AI algorithm is tested. Therefore, a degree of uncertainty is expected due to this nonuniformity.[92] A standardized approach for GT determination must be sought as it can be less user-dependent and capable of constructing generalizable algorithms.[49]

Future Directions

After a review of current trends in AI applications in the PET imaging of lymphoma, we aim to place a special focus on describing likely future directions in this field. The discussion related to the future directions is primarily aimed at depicting the potential clinical utility of AI in the management of lymphoma with ^{18}F-FDG-PET imaging.

From spatial domain to spatiotemporal realm—Current AI-based methods rarely use prior images in their models. This is counterintuitive for clinicians who consider prior images as one of the most important sources of information.[93, 94, 95] In clinical practice, temporal changes in a lesion may provide much more useful information than imaging features of a lesion in one study. As an example, conventional PET radiomics methods usually use a single-time-point for analysis, which does not take into account the interim change of the lesions throughout the treatment process.[96] However, "Delta-radiomics"[95] appraise the change in radiomic features during or after treatment to enhance information extraction.[95,97] Several studies have shown that CT-based Delta-radiomics can be used for the prediction of lung cancer, gastric cancer, and also the detection of side effects to radiation therapy.[94,98, 99, 100] Therefore, this method will be better suited to evaluate tumor response of treatment.[56] Creation of the Delta image permits identifying any posttreatment tumor transformations.[94,101] The difference between the baseline and interim parameters may be measured (ΔSUV, ΔMTV, ΔTLG, and ΔADC) for this purpose.[102] Delta radiomics

as a quantitative assessment can be used for the evaluation of the changes over time and for the prediction of treatment response earlier in the treatment course.[101] Time-interval changes in the features such as SUV_{max}, TMTV, MH, and D_{max} and the other features can be considered as the complementary important features that have not been considered before, especially for analysis of lymphoma data.

Using AI capabilities to generate and visualize Delta images will provide insight into the intralesional tumor heterogeneity, such as when a piece of a bulky tumor shrinks/improves while the other portion of the lesion grows. This capacity will be revolutionary for the detection and evaluation of tumor heterogeneity and heterogeneous response of tumor colonies to therapies. Future studies should also determine how to appropriately visualize the Delta image for better interpretability. Furthermore, the quantification of Delta images and identification of nonresponders at an earlier stage is a key direction AI-based algorithms can move toward.[103] This utility allows the optimization of treatment or biopsy of the non-responding lesion (or portion of a lesion) for new mutations. By identifying the nonresponder region of the tumor clinicians will also be able to use external beam radiation or percutaneous ablation sooner during lymphoma treatment. Recommendations for posttreatment lymphoma recurrence surveillance using AI-based PET imaging can be included in guidelines. Particularly in those with cancer remission, the Delta image could allow important insights on early and accurate detection of potential recurrence.

From data silos to large shared databases—Accessible, high-quality, and diverse imaging datasets are essential for accelerating the development of AI algorithms in lymphoma and successful transition of these technologies into routine clinical practice.[104] These repositories can bypass many barriers for researchers and diversify the patient population and their lymphoma subtypes, therefore, improving the generalizability of the algorithms. Especially when these datasets include multi-centric GT data and were generated by expert with varying levels of experience to recreate the heterogeneities that exist in real clinical practice rather than a controlled setting for biomedical research.

Currently, there is no centralized publicly available medical imaging data repository for lymphoma. There have, however, been minor initiatives to establish open access data sets. The Cancer Imaging Archive (TCIA) contains almost 31 million de-identified cancer medical images that are accessible to the public[105] which includes CT and ^{18}F-FDG PET studies of 155 DLBCL patients.[106] Additionally, the National Institutes of Health (NIH) has made available over 10,500 labeled CT imaging studies of 4400 different patients (DeepLesion dataset) with lung nodules, liver tumors, enlarged lymph nodes, and other critical findings throughout the body.[107] With the rapid growth of AI algorithms in lymphoma PET imaging, there is an increasing demand for institutional collaboration to enable the gathering and curation of both large data sets and labeled images necessary for the establishment of centralized yet diverse open access data repositories.

Summary

Almost 35 years after the first utilization of ^{18}F-FDG PET in lymphoma,[108,109] this modality has proven its value in a wide spectrum of management from diagnosis and

staging to treatment response assessment and prognostication. Alongside the significant improvement in treatment measures, from biological agents[110] to CAR-T cell therapy,[111] there have been substantial efforts to improve the quantitative aspect of ^{18}F-FDG PET to produce robust, reliable, and feasible metabolic imaging biomarkers. Establishment of PERCIST was a monumental event,[112] 2 years after universal acceptance of PET imaging for lymphoma response assessment by International Working Group (IWG).[113] ^{18}F-FDG PET was extremely successful in the elimination of "complete remission unconfirmed" (CRu) category in treatment response assessment. CRu was assigned to patients with a residual mass detected by CT after treatment that was unlikely to be malignant. Lugano classification,[1] a consensus document of the 11th International Conference on Malignant Lymphoma, reemphasized the role of ^{18}F-FDG-PET. Advancement of biological agents underscored the importance of the "tumor flare" phenomenon (pseudo-progression) and resulted in the refinement of Lugano classification by the introduction of immune-related criteria (IRC)[114] emphasizing the importance of clinical context in the process of image interpretation and quantification.

Treatment failure is the major problem in the management of patients with lymphoma and ^{18}F-FDG-PET has been providing valuable information to predict this event[115,116] and guide the treatment (response-adapted therapy).[117,118] Furthermore, the importance of biologic heterogeneity in treatment failure[119] motivated the molecular imaging community to detect and quantify this heterogeneity using Radiomics.[119]

In spite of all these achievements, the clinical adaptation of quantitative PET-based imaging biomarkers has been limited so far. Workflow integration barriers are one of the major contributing factors. Deep learning has the potential to make this process more efficient and more precise at the same time and this will open the door for all the subsequent utilization of PET-based imaging biomarkers. But the utility of AI is not limited to lesion detection and segmentation. Almost 20 years after wide utilization of ^{18}F-FDG-PET in lymphoma, we are experiencing a major transformation powered by AI affecting the entire imaging lifecycle: from scheduling and operational tasks [Beegle and colleagues' article "Artificial Intelligence and Positron Emission Tomography Imaging Workflow: Technologists' Perspective," in this issue], to image acquisition optimization [Muhammad Nasir Ullah and Craig S. Levin's article, "Application of Artificial Intelligence (AI) in Positron Tomography (PET) instrumentation," in this issue], enhancement of image reconstruction[120] and harmonization of the images[121] toward high-throughput imaging biomarkers[122] and multi-omics data integration for prediction and prognosis [Yousefirizi and colleagues' article, "Artificial Intelligence-Based Detection, Classification and Prediction/Prognosis in PET Imaging: Towards Radiophenomics," in this issue].[123]

Disclosure

This research was supported by the Intramural Research Program of the NIH, Clinical Center and NIDCR. The opinions expressed in this publication are the author's own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, or the United States government.

References

1. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non Hodgkin lymphoma: the Lugano classification J Clin Oncol, 32 (27) (2014), pp. 3059–3068 [PubMed: 25113753]
2. El-Galaly TC, Hutchings M, Mylam KJ, et al. Impact of18F-fluorodeoxyglucose positron emission tomography/computed tomography staging in newly diagnosed classical Hodgkin lymphoma: fewer cases with stage I disease and more with skeletal involvement Leuk Lymphoma, 55 (10) (2014), pp. 2349–2355 [PubMed: 24354682]
3. Hutchings M, Barrington SF PET/CT for therapy response assessment in lymphoma J Nucl Med, 50 (Suppl 1) (2009), pp. 21S–30S [PubMed: 19380407]
4. Baba S, Abe K, Isoda T, et al. Impact of FDG-PET/CT in the management of lymphoma Ann Nucl Med, 25 (10) (2011), pp. 701–716 [PubMed: 22037934]
5. Meignan M, Itti E, Gallamini A, et al. FDG PET/CT imaging as a biomarker in lymphoma Eur J Nucl Med Mol Imaging, 42 (4) (2015), pp. 623–633 [PubMed: 25573631]
6. Alavi A, Newberg AB, Souder E, et al. Quantitative analysis of PET and MRI data in normal aging and Alzheimer's disease: atrophy weighted total brain metabolism and absolute whole brain metabolism as reliable discriminators J Nucl Med, 34 (10) (1993), pp. 1681–1687 [PubMed: 8410281]
7. Saboury B, Salavati A, Brothers A, et al. FDG PET/CT in Crohn's disease: correlation of quantitative FDG PET/CT parameters with clinical and endoscopic surrogate markers of disease activity Eur J Nucl Med Mol Imaging, 41 (4) (2014), pp. 605–614 [PubMed: 24253895]
8. Saboury B, Parsons MA, Moghbel M, et al. Quantification of aging effects upon global knee inflammation by 18F-FDG-PET Nucl Med Commun, 37 (3) (2016), pp. 254–258 [PubMed: 26555103]
9. Basu S, Zaidi H, Salavati A, et al. FDG PET/CT methodology for evaluation of treatment response in lymphoma: from "graded visual analysis" and "semiquantitative SUVmax" to global disease burden assessment Eur J Nucl Med Mol Imaging, 41 (11) (2014), pp. 2158–2160 [PubMed: 24993455]
10. Basu S, Saboury B, Torigian DA, et al. Current evidence base of FDG-PET/CT imaging in the clinical management of malignant pleural mesothelioma: emerging significance of image segmentation and global disease assessment Mol Imaging Biol, 13 (5) (2011), pp. 801–811 [PubMed: 21136185]
11. Guo B, Tan X, Ke Q, et al. Prognostic value of baseline metabolic tumor volume and total lesion glycolysis in patients with lymphoma: a meta-analysis PLoS One, 14 (1) (2019), p. e0210224 [PubMed: 30625203]
12. Saboury B, Moghbel M, Basu S, et al. Modern Quantitative Techniques for PET/CT/MR Hybrid Imaging. In: Schaller B, ed. Molecular Imaging. IntechOpen; (2012)
13. Ceriani L, Milan L, Martelli M, et al. Metabolic heterogeneity on baseline 18FDG-PET/CT scan is a predictor of outcome in primary mediastinal B-cell lymphoma Blood, 132 (2) (2018), pp. 179–186 [PubMed: 29720487]
14. Ceriani L, Gritti G, Cascione L, et al. SAKK38/07 study: integration of baseline metabolic heterogeneity and metabolic tumor volume in DLBCL prognostic model. Blood Adv. 2020;4(6):1082–1092 Blood Adv, 4 (10) (2020), p. 2135 [PubMed: 32196557]
15. Cottereau A-S, Nioche C, Dirand A-S, et al. 18F-FDG PET dissemination features in diffuse large B-cell lymphoma are predictive of outcome J Nucl Med, 61 (1) (2020), pp. 40–45 [PubMed: 31201248]
16. Cottereau A-S, Meignan M, Nioche C, et al. New approaches in characterization of lesions dissemination in DLBCL patients on baseline PET/CT Cancers, 13 (16) (2021), 10.3390/cancers13163998
17. Foster B, Bagci U, Mansoor A, et al. A review on segmentation of positron emission tomography images Comput Biol Med, 50 (2014), pp. 76–96 [PubMed: 24845019]

18. Dewalle-Vignion A-S, Yeni N, Petyt G, et al. Evaluation of PET volume segmentation methods: comparisons with expert manual delineations Nucl Med Commun, 33 (1) (2012), pp. 34–42 [PubMed: 22044864]
19. Burggraaff CN, Rahman F, Kaßner I, et al. Optimizing workflows for fast and reliable metabolic tumor volume measurements in diffuse large B cell lymphoma Mol Imaging Biol, 22 (4) (2020), pp. 1102–1110 [PubMed: 31993925]
20. Zijlstra JM, Comans EF, van Lingen A, et al. FDG PET in lymphoma: the need for standardization of interpretation. An observer variation study Nucl Med Commun, 28 (10) (2007), pp. 798–803 [PubMed: 17728610]
21. Frood R, Burton C, Tsoumpas C, et al. Baseline PET/CT imaging parameters for prediction of treatment outcome in Hodgkin and diffuse large B cell lymphoma: a systematic review Eur J Nucl Med Mol Imaging (2021), 10.1007/s00259-021-05233-2
22. Annunziata S, Pelliccioni A, Hohaus S, et al. The prognostic role of end-of-treatment FDG-PET/CT in diffuse large B cell lymphoma: a pilot study application of neural networks to predict time-to-event Ann Nucl Med, 35 (1) (2021), pp. 102–110 [PubMed: 33094420]
23. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies Ann Intern Med, 155 (8) (2011), pp. 529–536 [PubMed: 22007046]
24. Bi L, Kim J, Kumar A, et al. Automatic detection and classification of regions of FDG uptake in whole-body PET-CT lymphoma studies Comput Med Imaging Graph, 60 (2017), pp. 3–10 [PubMed: 27955798]
25. Barrington SF, Meignan M Time to prepare for risk adaptation in lymphoma by standardizing measurement of metabolic tumor burden J Nucl Med, 60 (8) (2019), pp. 1096–1102 [PubMed: 30954945]
26. Berkowitz A, Basu S, Srinivas S, et al. Determination of whole-body metabolic burden as a quantitative measure of disease activity in lymphoma: a novel approach with fluorodeoxyglucose-PET Nucl Med Commun, 29 (6) (2008), pp. 521–526 [PubMed: 18458598]
27. Akhtari M, Milgrom SA, Pinnix CC, et al. Reclassifying patients with early-stage Hodgkin lymphoma based on functional radiographic markers at presentation Blood, 131 (1) (2018), pp. 84–94 [PubMed: 29038339]
28. Kostakoglu L, Goldsmith SJ 18F-FDG PET evaluation of the response to therapy for lymphoma and for breast, lung, and colorectal carcinoma J Nucl Med, 44 (2) (2003), pp. 224–239 [PubMed: 12571214]
29. Gallamini A PET scan in Hodgkin lymphoma: role in diagnosis, prognosis, and treatment Springer, New York City, NY, USA (2016)
30. Punwani S, Taylor SA, Saad ZZ, et al. Diffusion-weighted MRI of lymphoma: prognostic utility and implications for PET/MRI? Eur J Nucl Med Mol Imaging, 40 (3) (2013), pp. 373–385 [PubMed: 23197155]
31. Gull S, Akbar S Artificial intelligence in brain tumor detection through MRI Scans Artif Intelligence Internet Things (2021), pp. 241–276, 10.1201/9781003097204-10
32. Yousefirizi F, Jha AK, Brosch-Lenz J, et al. Towards high-throughput AI-based segmentation in oncological PET imaging. arXiv [physics.med-ph] Available at: http://arxiv.org/abs/2107.13661 (2021) Accessed September 10, 2021
33. Taghanaki SA, Abhishek K, Cohen JP, et al. Deep semantic segmentation of natural and medical images: a review Artif Intelligence Rev, 54 (1) (2021), pp. 137–178
34. Hirata K, Manabe O, Magota K, et al. A preliminary study to use SUVmax of FDG PET-CT as an identifier of lesion for artificial intelligence Front Med, 8 (2021), p. 647562
35. Spatial and temporal image registration Yankeelov TE, Pickens DR, Price RR (Eds.), Quantitative MRI in cancer, CRC Press, Boca Raton, FL (2011), pp. 256–269
36. Jiao J, Searle GE, Tziortzi AC, et al. Spatio-temporal pharmacokinetic model based registration of 4D PET neuroimaging data. Neuroimage. 2014;84:225–235. [PubMed: 23994455]
37. Pereira G Deep Learning techniques for the evaluation of response to treatment in Hogdkin Lymphoma Available at: https://estudogeral.uc.pt/handle/10316/86276 (2018) Accessed September 10, 2021

38. Zhou Z, Jain P, Lu Y, et al. Computer-aided detection of mantle cell lymphoma on 18F-FDG PET/CT using a deep learning convolutional neural network Am J Nucl Med Mol Imaging, 11 (4) (2021), p. 260 [PubMed: 34513279]
39. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation Ann Intern Med, 169 (7) (2018), pp. 467–473 [PubMed: 30178033]
40. Mayerhoefer ME, Umutlu L, Schöder H Functional imaging using radiomic features in assessment of lymphoma Methods, 188 (2021), pp. 105–111 [PubMed: 32634555]
41. Sheng VS, Provost F, Ipeirotis PG. Get another label? improving data quality and data mining using multiple, noisy labelers. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '08. Association for Computing Machinery; Las Vegas, Nevada, USA: 8 24–27, 2008:614–622.
42. Park SH, Han K Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction Radiology, 286 (3) (2018), pp. 800–809 [PubMed: 29309734]
43. Pinochet P, Eude F, Becker S, et al. Evaluation of an automatic classification algorithm using convolutional neural networks in oncological positron emission tomography Front Med, 8 (2021), p. 628179
44. Sadik M, López-Urdaneta J, Ulén J, et al. Artificial intelligence could alert for focal skeleton/bone marrow uptake in Hodgkin's lymphoma patients staged with FDG-PET/CT Sci Rep, 11 (1) (2021), p. 10382 [PubMed: 34001922]
45. Guo R, Hu X, Song H, et al. Weakly supervised deep learning for determining the prognostic value of 18F-FDG PET/CT in extranodal natural killer/T cell lymphoma, nasal type Eur J Nucl Med Mol Imaging (2021), 10.1007/s00259-021-05232-3
46. Yuan C, Zhang M, Huang X, et al. Diffuse large B-cell lymphoma segmentation in PET-CT images via hybrid learning for feature fusion Med Phys (2021), 10.1002/mp.14847 mp.14847
47. Blanc-Durand P, Jégou S, Kanoun S, et al. Fully automatic segmentation of diffuse large B cell lymphoma lesions on 3D FDG-PET/CT for total metabolic tumour volume prediction using a convolutional neural network Eur J Nucl Med Mol Imaging, 48 (5) (2021), pp. 1362–1370 [PubMed: 33097974]
48. Weisman AJ, Kim J, Lee I, et al. Automated quantification of baseline imaging PET metrics on FDG PET/CT images of pediatric Hodgkin lymphoma patients EJNMMI Phys, 7 (1) (2020), p. 76 [PubMed: 33315178]
49. Weisman AJ, Kieler MW, Perlman SB, et al. Convolutional neural networks for automated PET/CT detection of diseased lymph node burden in patients with lymphoma Radiol Artif Intell, 2 (5) (2020), p. e200016 [PubMed: 33937842]
50. Sibille L, Seifert R, Avramovic N, et al. 18F-FDG PET/CT uptake classification in lymphoma and lung cancer by using deep convolutional neural networks Radiology, 294 (2) (2020), pp. 445–452 [PubMed: 31821122]
51. Li H, Jiang H, Li S, et al. DenseX-Net: an end-to-end model for lymphoma segmentation in whole-body PET/CT Images IEEE Access, 8 (2020), pp. 8004–8018
52. Sadik M, Lind E, Polymeri E, et al. Automated quantification of reference levels in liver and mediastinal blood pool for the Deauville therapy response classification using FDG-PET/CT in Hodgkin and non-Hodgkin lymphomas Clin Physiol Funct Imaging, 39 (1) (2019), pp. 78–84 [PubMed: 30284376]
53. Goodfellow I, Bengio Y, Courville A Deep learning MIT Press, Cambridge, MA, USA (2016)
54. Lippi M, Gianotti S, Fama A, et al. Texture analysis and multiple-instance learning for the classification of malignant lymphomas Comput Methods Programs Biomed, 185 (2020), p. 105153 [PubMed: 31678792]
55. Mayerhoefer ME, Riedl CC, Kumar A, et al. Radiomic features of glucose metabolism enable prediction of outcome in mantle cell lymphoma Eur J Nucl Med Mol Imaging, 46 (13) (2019), pp. 2760–2769 [PubMed: 31286200]
56. Hu H, Decazes P, Vera P, et al. Detection and segmentation of lymphomas in 3D PET images via clustering with entropy-based optimization strategy Int J Comput Assist Radiol Surg, 14 (10) (2019), pp. 1715–1724 [PubMed: 31401714]

57. Yu Y, Decazes P, Lapuyade-Lahorgue J, et al. Semi-automatic lymphoma detection and segmentation using fully conditional random fields Comput Med Imaging Graph, 70 (2018), pp. 1–7 [PubMed: 30253305]
58. Grossiord É, Talbot H, Passat N, Meignan M, Najman L. Automated 3D lymphoma lesion segmentation from PET/CT characteristics. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). Melbourne, VIC, Australia; 18–21 April 2017:174–178.
59. Desbordes P, Petitjean C, Ruan S Segmentation of lymphoma tumor in PET images using cellular automata: a preliminary study IRBM, 37 (1) (2016), pp. 3–10
60. Lartizien C, Rogez M, Niaf E, et al. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information IEEE J Biomed Health Inform, 18 (3) (2014), pp. 946–955 [PubMed: 24081876]
61. Weisman A, Kim J, Lee I, et al. Automated deep learning-based quantification of baseline imaging PET metrics on FDG PET/CT images of pediatric lymphoma patients J Nucl Med, 61 (Suppl 1) (2020), p. 506
62. Puttagunta M, Ravi S Medical image analysis based on deep learning approach Multimed Tools Appl (2021), pp. 1–34
63. Kostakoglu L, Chauvie S PET-derived quantitative metrics for response and prognosis in lymphoma PET Clin, 14 (3) (2019), pp. 317–329 [PubMed: 31084772]
64. Lucignani SUV G and segmentation: pressing challenges in tumour assessment and treatment Eur J Nucl Med Mol Imaging, 36 (4) (2009), pp. 715–720 [PubMed: 19205697]
65. Yousefirizi F, Jha AK, Brosch-Lenz J, et al. Toward high-throughput artificial intelligence-based segmentation in oncological PET imaging PET Clin, 16 (4) (2021), pp. 577–596 [PubMed: 34537131]
66. Vercellino L, Cottereau A-S, Casasnovas O, et al. High total metabolic tumor volume at baseline predicts survival independent of response to therapy Blood, 135 (16) (2020), pp. 1396–1405 [PubMed: 31978225]
67. Weisman AJ, Kieler MW, Perlman S, et al. Comparison of 11 automated PET segmentation methods in lymphoma Phys Med Biol, 65 (23) (2020), p. 235019 [PubMed: 32906088]
68. Rahim MK, Kim SE, So H, et al. Recent trends in PET image interpretations using volumetric and texture-based quantification methods in nuclear oncology Nucl Med Mol Imaging, 48 (1) (2014), pp. 1–15 [PubMed: 24900133]
69. Rizzo A, Triumbari EKA, Gatta R, et al. The role of 18F-FDG PET/CT radiomics in lymphoma Clin Translational Imaging (2021), 10.1007/s40336-021-00451-y
70. Starmans MPA, van der Voort SR, Castillo Tovar JM, et al. Chapter 18 - radiomics: Data mining using quantitative medical image features Zhou SK, Rueckert D, Fichtinger G (Eds.), Handbook of medical image computing and computer assisted intervention, Academic Press, Cambridge, MA, US (2020), pp. 429–456
71. Hatt M, Le Rest CC, Tixier F, et al. Radiomics: data are also images J Nucl Med, 60 (Suppl 2) (2019), pp. 38S–44S [PubMed: 31481588]
72. Blanc-Durand P, Van Der Gucht A, Schaefer N, et al. Automatic lesion detection and segmentation of 18FET PET in gliomas : a full 3D U-Net convolutional neural network study J Nucl Med, 59 (Suppl 1) (2018), p. 330
73. Klyuzhin I, Xu Y, Harsini S, et al. Unsupervised background removal by dual-modality PET/CT guidance: application to PSMA imaging of metastases J Nucl Med, 62 (Suppl 1) (2021), p. 36
74. Ou X, Zhang J, Wang J, et al. Radiomics based on 18 F-FDG PET/CT could differentiate breast carcinoma from breast lymphoma using machine-learning approach: a preliminary study Cancer Med, 9 (2) (2020), pp. 496–506 [PubMed: 31769230]
75. Xu H, Guo W, Cui X, et al. Three-dimensional texture analysis based on PET/CT images to distinguish hepatocellular carcinoma and hepatic lymphoma Front Oncol, 9 (2019), p. 844 [PubMed: 31552173]
76. Ou X, Wang J, Zhou R, et al. Ability of 18F-FDG PET/CT radiomic features to distinguish breast carcinoma from breast lymphoma Contrast Media Mol Imaging, 2019 (2019), p. 4507694 [PubMed: 30930700]

77. Aide N, Talbot M, Fruchart C, et al. Diagnostic and prognostic value of baseline FDG PET/CT skeletal textural features in diffuse large B cell lymphoma Eur J Nucl Med Mol Imaging, 45 (5) (2018), pp. 699–711 [PubMed: 29214417]

78. Rodríguez Taroco MG, Cuña EG, Pages C, et al. Prognostic value of imaging markers from 18FDG-PET/CT in paediatric patients with Hodgkin lymphoma Nucl Med Commun, 42 (3) (2021), pp. 306–314 [PubMed: 33306628]

79. Eertink JJ, van de Brug T, Wiegers SE, et al. 18F-FDG PET baseline radiomics features improve the prediction of treatment outcome in diffuse large B-cell lymphoma Eur J Nucl Med Mol Imaging (2021), 10.1007/s00259-021-05480-3

80. Wang H, Zhao S, Li L, et al. Development and validation of an 18F-FDG PET radiomic model for prognosis prediction in patients with nasal-type extranodal natural killer/T cell lymphoma Eur Radiol, 30 (10) (2020), pp. 5578–5587 [PubMed: 32435928]

81. Sun Y, Qiao X, Jiang C, et al. Texture analysis improves the value of pretreatment 18F-FDG PET/CT in predicting interim response of primary gastrointestinal diffuse large B-cell lymphoma Contrast Media Mol Imaging, 2020 (2020), p. 2981585 [PubMed: 32922221]

82. Aide N, Fruchart C, Nganoa C, et al. Baseline 18F-FDG PET radiomic features as predictors of 2-year event-free survival in diffuse large B cell lymphomas treated with immunochemotherapy Eur Radiol, 30 (8) (2020), pp. 4623–4632 [PubMed: 32248365]

83. Wu J, Lian C, Ruan S, et al. Treatment outcome prediction for cancer patients based on radiomics and belief function theory IEEE Trans Radiat Plasma Med Sci, 3 (2) (2019), pp. 216–224 [PubMed: 31903444]

84. Tatsumi M, Isohashi K, Matsunaga K, et al. Volumetric and texture analysis on FDG PET in evaluating and predicting treatment response and recurrence after chemotherapy in follicular lymphoma Int J Clin Oncol, 24 (10) (2019), pp. 1292–1300 [PubMed: 31165310]

85. Lue K-H, Wu Y-F, Liu S-H, et al. Prognostic value of pretreatment radiomic features of 18F-FDG PET in patients with Hodgkin lymphoma Clin Nucl Med, 44 (10) (2019), pp. e559–e565 [PubMed: 31306204]

86. Lue K-H, Wu Y-F, Liu S-H, et al. Intratumor heterogeneity assessed by 18F-FDG PET/CT predicts treatment response and survival outcomes in patients with Hodgkin lymphoma Acad Radiol, 27 (8) (2020), pp. e183–e192 [PubMed: 31761665]

87. Zhou Y, Ma X-L, Pu L-T, et al. Prediction of Overall survival and progression-free survival by the 18F-FDG PET/CT radiomic features in patients with primary gastric diffuse large B-cell lymphoma Contrast Media Mol Imaging, 2019 (2019), p. 5963607 [PubMed: 31777473]

88. Milgrom SA, Elhalawani H, Lee J, et al. A PET radiomics model to predict refractory mediastinal Hodgkin lymphoma Sci Rep, 9 (1) (2019), 10.1038/s41598-018-37197-z

89. Wang M, Xu H, Xiao L, et al. Prognostic value of functional parameters of 18F-FDG-PET images in patients with primary renal/adrenal lymphoma Contrast Media Mol Imaging, 2019 (2019), p. 2641627 [PubMed: 31427906]

90. Parvez A, Tau N, Hussey D, et al. Erratum to: 18F-FDG PET/CT metabolic tumor parameters and radiomics features in aggressive non Hodgkin's lymphoma as predictors of treatment outcome and survival Ann Nucl Med, 32 (6) (2018), pp. 410–416 [PubMed: 29754276]

91. Ben Bouallègue F, Tabaa YA, Kafrouni M, et al. Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas Med Phys, 44 (9) (2017), pp. 4608–4619 [PubMed: 28513853]

92. Pfaehler E, Burggraaff C, Kramer G, et al. PET segmentation of bulky tumors: strategies and workflows to improve inter-observer variability PLoS One, 15 (3) (2020), p. e0230901 [PubMed: 32226030]

93. Noortman WA, Vriens D, Slump CH, et al. Adding the temporal domain to PET radiomic features PLoS One, 15 (9) (2020), p. e0239438 [PubMed: 32966313]

94. Fave X, Zhang L, Yang J, et al. Delta-radiomics features for the prediction of patient outcomes in non–small cell lung cancer Sci Rep, 7 (1) (2017), pp. 1–11 [PubMed: 28127051]

95. Carvalho S, Leijenaar RTH, Troost EGC, et al. Early variation of FDG-PET radiomics features in NSCLC is related to overall survival - the "delta radiomics" concept Radiother Oncol, 118 (Suppl 1) (2016), pp. S20–S21

96. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to radiomics J Nucl Med, 61 (4) (2020), pp. 488–495 [PubMed: 32060219]

97. Alahmari SS, Cherezov D, Goldgof D, et al. Delta radiomics improves pulmonary nodule malignancy prediction in lung cancer screening IEEE Access, 6 (2018), pp. 77796–77806 [PubMed: 30607311]

98. Wang L, Gao Z, Li C, et al. Computed tomography--based delta-radiomics analysis for discriminating radiation pneumonitis in patients with esophageal cancer after radiation therapy Int J Radiat Oncol Biol Phys (2021) Available at: https://www.sciencedirect.com/science/article/pii/S0360301621004752?casa_token=klpUfNmgEhEAAAA:ZOAShjEEgzDXK6JxuvCpIWKcps-6o7x51hP4a952C9kqQMbH7zXrqgkjIhumgcLoWkrVJu8

99. Mazzei MA, Di Giacomo L, Bagnacci G, et al. Delta-radiomics and response to neoadjuvant treatment in locally advanced gastric cancer—a multicenter study of GIRCG (Italian Research Group for Gastric Cancer) Quantitative Imaging Med Surg, 11 (6) (2021), pp. 2376–2387

100. Liu Y, Shi H, Huang S, et al. Early prediction of acute xerostomia during radiation therapy for nasopharyngeal cancer based on delta radiomics from CT images Quant Imaging Med Surg, 9 (7) (2019), pp. 1288–1302 [PubMed: 31448214]

101. Nasief H, Zheng C, Schott D, et al. A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer NPJ Precis Oncol, 3 (2019), p. 25 [PubMed: 31602401]

102. Fave X, Zhang L, Yang J, et al. Using pretreatment radiomics and delta-radiomics features to predict non–small cell lung cancer patient outcomes Int J Radiat Oncol Biol Phys, 98 (1) (2017), p. 249

103. Bera K, Velcheti V, Madabhushi A Novel quantitative imaging for predicting response to therapy: techniques and clinical applications Am Soc Clin Oncol Educ Book, 38 (2018), pp. 1008–1018 [PubMed: 30231314]

104. Kohli MD, Summers RM, Geis JR Medical image data and datasets in the era of machine learning-whitepaper from the 2016 C-MIMI meeting dataset session J Digit Imaging, 30 (4) (2017), pp. 392–399 [PubMed: 28516233]

105. Prior F, Smith K, Sharma A, et al. The public cancer radiology imaging collections of the cancer imaging archive Sci Data, 4 (2017), p. 170124 [PubMed: 28925987]

106. Dose-adjusted EPOCH-R compared with R-CHOP as frontline therapy for diffuse large B-cell lymphoma (CALGB50303) - the cancer imaging archive (TCIA) public access - cancer imaging archive wiki Available at: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70225094 Accessed October 1, 2021

107. Yan K, Wang X, Lu L, et al. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning J Med Imaging (Bellingham), 5 (3) (2018), p. 036501 [PubMed: 30035154]

108. Kiyosawa M, Ohmura M, Mizuno K, et al. [18F-FDG positron emission tomography in orbital lymphoid tumor] Nihon Ganka Gakkai Zasshi, 89 (12) (1985), pp. 1329–1333 [PubMed: 3879574]

109. Kuwabara Y, Ichiya Y, Otsuka M, et al. High [18F]FDG uptake in primary cerebral lymphoma: a PET study J Comput Assist Tomogr, 12 (1) (1988), pp. 47–48 [PubMed: 3257221]

110. Coiffier B, Lepage E, Briere J, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large B-cell lymphoma N Engl J Med, 346 (4) (2002), pp. 235–242 [PubMed: 11807147]

111. Schuster SJ, Bishop MR, Tam CS, et al. Tisagenlecleucel in adult relapsed or refractory diffuse large B-cell lymphoma N Engl J Med, 380 (1) (2019), pp. 45–56 [PubMed: 30501490]

112. Wahl RL, Jacene H, Kasamon Y, et al. From RECIST to PERCIST: evolving Considerations for PET response criteria in solid tumors J Nucl Med, 50 (Suppl 1) (2009), pp. 122S–150S [PubMed: 19403881]

113. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma J Clin Oncol, 25 (5) (2007), pp. 579–586 [PubMed: 17242396]

114. Cheson BD, Ansell S, Schwartz L, et al. Refinement of the Lugano Classification lymphoma response criteria in the era of immunomodulatory therapy Blood, 128 (21) (2016), pp. 2489–2496 [PubMed: 27574190]
115. Hutchings M, Loft A, Hansen M, et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma Blood, 107 (1) (2006), pp. 52–59 [PubMed: 16150944]
116. Burggraaff CN, de Jong A, Hoekstra OS, et al. Predictive value of interim positron emission tomography in diffuse large B-cell lymphoma: a systematic review and meta-analysis Eur J Nucl Med Mol Imaging, 46 (1) (2019), pp. 65–79 [PubMed: 30141066]
117. André MPE, Girinsky T, Federico M, et al. Early positron emission tomography response-adapted treatment in stage I and II Hodgkin lymphoma: final results of the randomized EORTC/LYSA/FIL H10 trial J Clin Oncol, 35 (16) (2017), pp. 1786–1794 [PubMed: 28291393]
118. Borchmann P, Goergen H, Kobe C, et al. PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group Lancet, 390 (10114) (2017), pp. 2790–2802 [PubMed: 29061295]
119. Sehn LH, Gascoyne RD Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity Blood, 125 (1) (2015), pp. 22–32 [PubMed: 25499448]
120. 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, 16 (4) (2021 10), pp. 533–542, 10.1016/j.cpet.2021.06.004 [PubMed: 34537129]
121. Liu J, Malekzadeh M, Mirian N, Song TA, Liu C, Dutta J Artificial Intelligence-Based Image Enhancement in PET Imaging: Noise Reduction and Resolution Enhancement PET Clin, 16 (4) (2021 10), pp. 553–576, 10.1016/j.cpet.2021.06.005 [PubMed: 34537130]
122. Yousefirizi F, Jha AK, Brosch-Lenz J, Saboury B, Rahmim A Toward High-Throughput Artificial Intelligence-Based Segmentation in Oncological PET Imaging PET Clin, 16 (4) (2021 10), pp. 577–596, 10.1016/j.cpet.2021.06.001 [PubMed: 34537131]
123. 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(4):493–511. [PubMed: 34537127]

Key points

- One of the most serious issues in the management of lymphoma patients is treatment failure.
- Accurate quantification of tumor burden using ^{18}F-FDG-PET is an important method for therapy response assessment and prediction.
- Artificial Intelligence (AI)-based PET approaches could make this process more efficient, precise, and pave the way for future PET-based imaging biomarker applications.
- In addition to streamlining the workflow, AI can enable segmentation and radiomic analysis to acquire prognostic information regarding therapy augmentation, remission planning, and recurrence prediction.

Box 1

Abbreviations

AI	Artificial Intelligence	ML	Metabolic Heterogeneity
CNN	Convolutional Neural Network	MTV	Metabolic Tumor Volume
DL	Deep Learning	NHL	Non-Hodgkin Lymphoma
DLBCL	Diffuse Large B-cell Lymphoma	NM	Nuclear Medicine
FP	False Positive	OS	Overall Survival
FN	False Negative	PFS	Progression-Free Survival
^{18}F-FDG	^{18}F-fluorodeoxyglucose	RF	Random Forest
HL	Hodgkin Lymphoma	SVM	Support Vector Machine
HiNA	High Normal Activity	TLG	Total Lesion Glycolysis
ML	Machine Learning	TMTV	Total Metabolic Tumor Volume

Clinics care points

- In order to efficiently extract valuable information about tumor biology from ^{18}F-FDG-PET, we need to move beyond SUV measurement. The first step in this path is detection and delineation of hypermetabolic lesions.
- AI based methods have the potential to provide clinicians with a high throughput platform to perform these steps efficiently and accurately.
- Clinicians should be aware of pearls and pitfalls of AI algorithms. Deep learning is very efficient when utilized in the correct setting and could be prone to bias and aberrant performance if used out of scope of training and testing. It is ultimately the responsibility of physicians and the healthcare system to verify the trustworthiness and reliability of AI as Medical Devices (AIMDs).
- Despite their significant advances, PET-based AI applications have had limited clinical implementation. Immaturity of PACS architecture is among the important reasons. AI orchestrators will play an important role in future of imaging workflow.

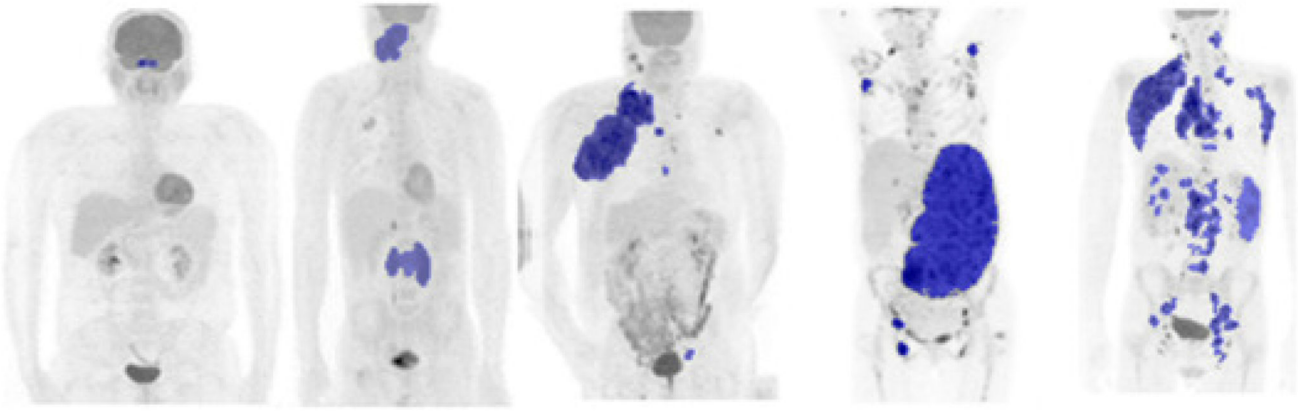

Fig. 1.

Examples of different sizes and distributions of tumor in 5 patients with diffuse large B-cell lymphoma[25].

(*From* Barrington SF, Meignan M. Time to Prepare for Risk Adaptation in Lymphoma by Standardizing Measurement of Metabolic Tumor Burden. *J Nucl Med*. 2019;60(8):1096–1102: under Open Access Creative Commons License https://creativecommons.org/licenses/by/4.0/)

1. Download : Download high-res image (173KB)
2. Download : Download full-size image

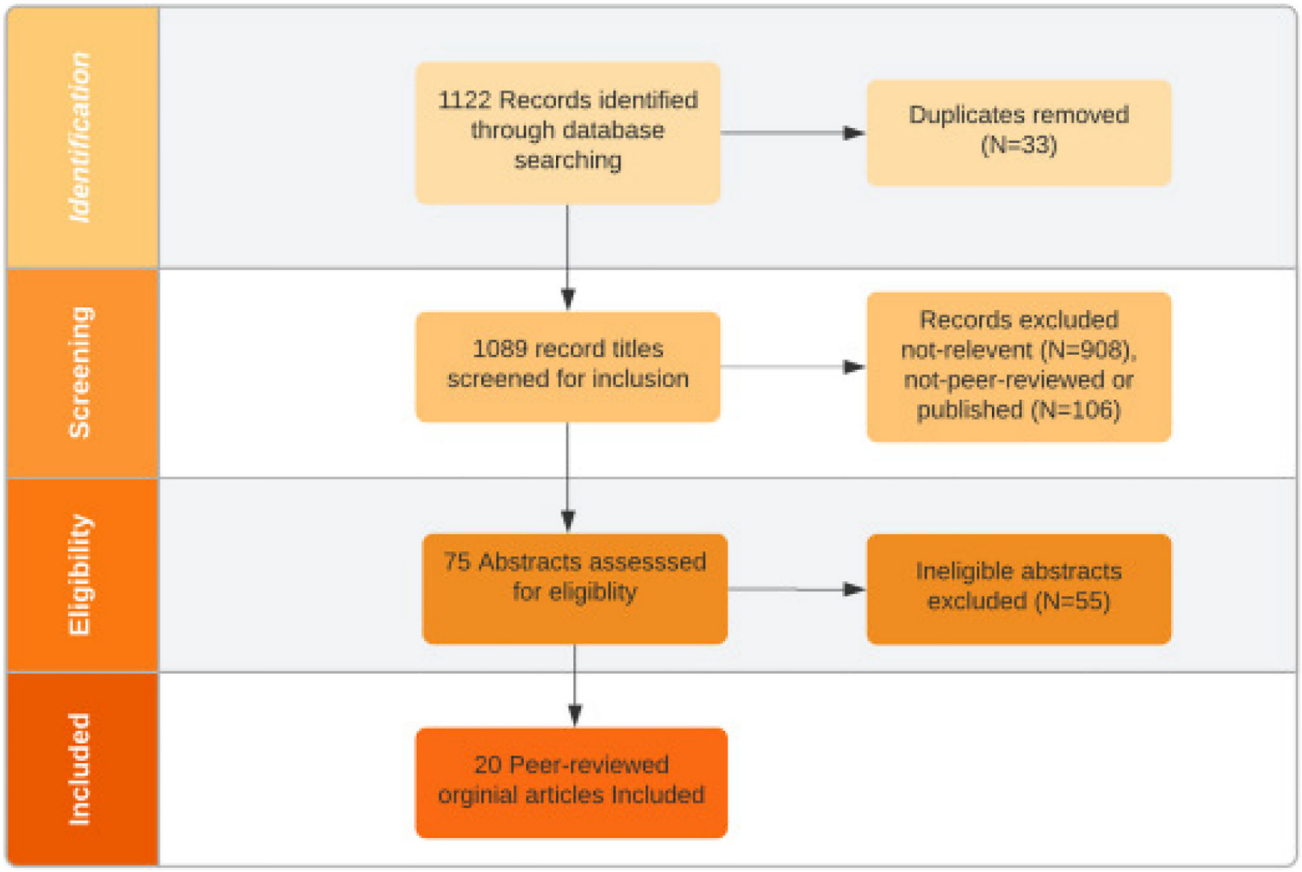

Fig. 2.

Demonstrates the summary of the literature search strategies and the results at each stage.

1. Download : Download high-res image (331KB)
2. Download : Download full-size image

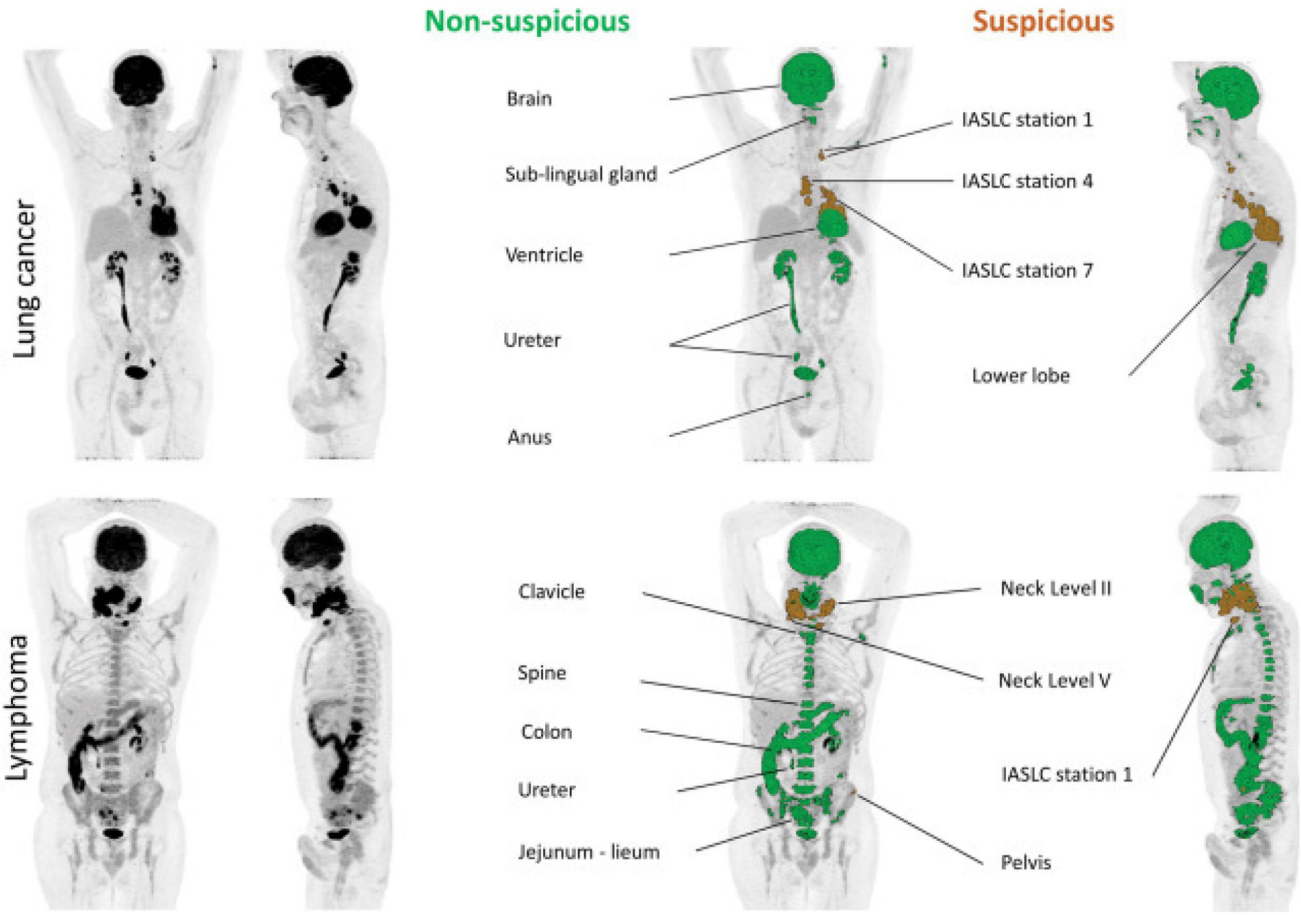

Fig. 3.

Maximum intensity projection ^{18}F-FDG PET/CT images were processed in 2 patients using the constructed CNN. The test data consists of patients with both lung cancer and lymphoma; the detected lesions are color coded accordingly. IASLC is the abbreviation for the International Association for the Study of Lung Cancer.

(*From* Sibille L, Seifert R, Avramovic N, et al. 18F-FDG PET/CT Uptake Classification in Lymphoma and Lung Cancer by Using Deep Convolutional Neural Networks. Radiology. 2020;294(2):445–452; with permission.)

1. Download : Download high-res image (584KB)
2. Download : Download full-size image

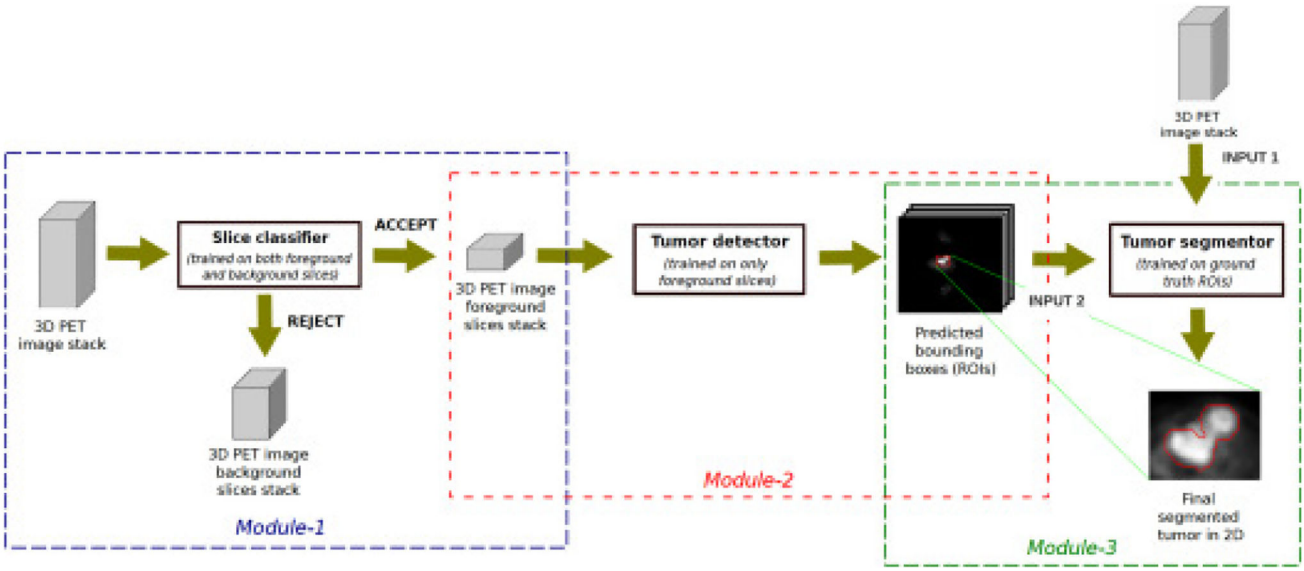

Fig. 4.

Schematic of a proposed cascaded model for PET tumor segmentation; Module 1, classifies the axial slices to suspicious and non-suspicious ones; Module 2, detects the lymphoma lesions in axial slices that are a candidate by Module 1. In Module 3, the 3D PET image and detection results are given to the tumor segmentation algorithm to segment the lesions inside the bounding boxes provided by Module 2.

1. Download : Download high-res image (278KB)
2. Download : Download full-size image

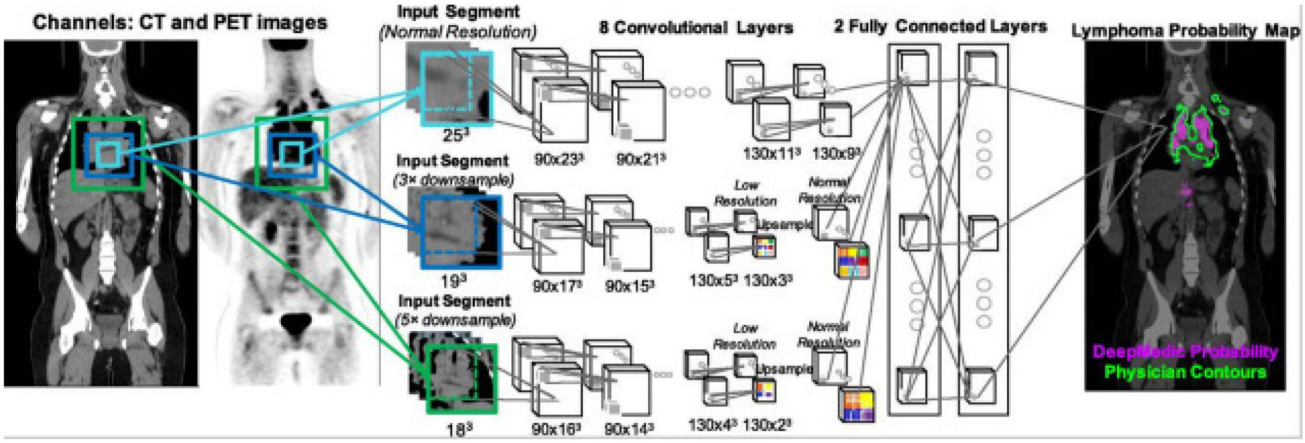

Fig. 5.

CT and coronal PET multicenter images are input to 3 segment layers, there are then 8 convolution layers and two fully connected layers that subsequently generate a probability map for lymphoma lesions as shown in purple.

(From Weisman AJ, Kim J, Lee I, et al. Automated quantification of baseline imaging PET metrics on FDG PET/CT images of pediatric Hodgkin lymphoma patients. *EJNMMI Phys.* 2020;7(1):76: under Open Access Creative Commons License http://creativecommons.org/licenses/by/4.0/.)

1. Download : Download high-res image (533KB)
2. Download : Download full-size image

Table 1.

Summary of characteristics of selected literature

Author (Year)	Tasks Performed	Task Specific Input/Out	FoM	Details (Model-Related)	Details (Ground Truth, Sample Size)
CNN Models					
Pinochet et al,[43] 2021	Classification (Radiophenomics)	Input: 2D; WB; axial/sagittal/coronal ^{18}F-FDG PET slice Output: slice-level 3-category classification (Benign, Malignant, Equivocal lymph nodes)	AUC = 0.62	Evaluate PET Assisted Reporting System (PARS-PET) by Siemens on DLBCL patients CNN model: PET Assisted Reporting System (PARS)	GT: 2 NM physicians segmented DLBCL lesions Testing: 119 patients (research cohort) + 430 patients (routine cohort)
	Segmentation	Input 1: 2D; WB; axial/sagittal/coronal ^{18}F-FDG PET slices Input 2: detection map provided by PARS prototype software Output: 2D; segmented lesions with borders masked on PET slice	DSC = 0.65 (research cohort) DSC = 0.48 (routine cohort) TMTV ICC = 0.68 (research cohort) TMTV ICC = 0.61 (routine cohort)		
	Statistical Classification (Prediction/Prognosis)	Input: TMTV value from WB PET Image (Method: TMTV thresholding) Output: Prognostication (PFS, OS)	OS Hazard ratio = 2.4 PFS Hazard ratio = 2.1		
Sadik et al,[44] 2021	Classification (Radiophenomics)	Input: 2D; WB; sagittal/coronal/axial; ^{18}F-FDG PET slice and CT slice (2 channel input) Output: Pt-level 4-class classification [high vs low diffuse bone marrow uptake] × [presence vs absence of focal lesion])	PA = 0.85 Kappa = 0.41	Highlight foci of skeletal and bone marrow uptake in Hodgkin's Lymphoma patients CNN model: based on RECOMIA prototype	GT: 10 independent experienced NM clinicians classified lesions Training: 156 patients Testing: 49 patients
Guo et al,[45] 2021	Characterization (deepRadiomics)	Input: Manually segmented lesions (3D; axial; 1 channel: Rank 3 Tensor [combined ^{18}F-FDG PET and CT]) Output: 16 × 8 feature maps, total of 128 features	AUC = 0.88 (for PSI) PSI-based PFS prediction: Spec = 0.80, Sens = 0.83, Accuracy = 0.85	Extraction of feature maps surrogates for prognosis prediction in nasal ENKTL. Proposes PSI be a predictor of PFS; PSI is the ratio of the PPV to NPV Model: Weakly supervised deep learning (WSDL) based on Residual Network-18 (ResNet-18) and PNU classifier.	GT: 1 NM physician (15 y experience) segmented nasal ENKTL lesions Training sample: 64 patients Testing: 20 patients
	Statistical Classification (Radiophenomics)	Input: Prediction similarity index (PSI) derived from image features Output: Relapsed vs nonrelapsed classes for ENKTL			
Yuan et al,[46] 2021	Detection	Input: 2D; axial; neck/chest/abdomen; ^{18}F-FDG PET slice and CT slices (2 channel input) Output: 2D; axial; detection map with lesions in rectangular boxes	Sens (chest) = 83.2%, Spec (chest) = 99.75%, Accuracy = 99.5%	Hybrid Learning for feature fusion of DLBCL Segmentation Hybrid CNN models can create feature fusion maps and quantify the spatial contributions of each modality. PET and CT image feature-based hybrid learning CNN model architecture	GT: 1 physician manually segmented DLBCL lesions Training and Validation: Cross-validation using a dataset with total of 1242 PET-CT slice pairs from 45 PET-CT samples

Author (Year)	Tasks Performed	Task Specific Input/Out	FoM	Details (Model-Related)	Details (Ground Truth, Sample Size)
	Segmentation	Input 1: 2D; axial; neck/chest/abdomen; ^{18}F-FDG PET slice and CT slices (2 channel input) Input 2: Detection map results Output: lesions border segmentation map	DSC = 0.73, MHD = 4.38 mm		
Zhou et al,[38] 2021	Detection	Input: 2D; axial/coronal; WB ^{18}F-FDG PET/CT (1 channel) or ^{18}F-FDG PET alone Output: Map of detected mantle cell lymphoma	Sens = 0.88 FP/patient = 15 For outside-institute patients Sens = 0.84 FP/patient = 14	Xception-based U-Net Localized lesions on PET/CT and labeled each pixel as MCL or not MCL. High FPs/patient needs to be corrected through physicians' inspection	GT: 3 NM physicians each with more than 10 y of experience identified and contoured MCL lesions Training: 110 patients Validation: 5-fold cross-validation Testing: 32 outside-institute patients
	End-to-end Segmentation	Input: 3D; coronal; WB; ^{18}F-FDG PET and 3D CT 2 separate channels Output: Mask of segmented lesions with calculated TMTV on 18F-FDG PET/CT	JSC = 0.60, DSC = 0.73 Predicted TMTV R = 0.88, 0.82 in first cohort, second cohort, respectively	Fully automatic segmentation of DLBCL lesions for total MTV prediction – 3D ^{18}FDG-PET/CT	GT: masks were manually obtained with 41% SUV_{max} adaptive thresholding. TMTV protocol from LIFEx used for VOI semiautomatically segmented. 2 experienced physicians reviewed clustering results and remove physiologic uptakes Training: 639 patients Validation: 5-fold cross-validation Testing: 94 patients
Weisman et al,[48] 2020	End-to-end Segmentation	Input: 3D; coronal; WB; ^{18}F-FDG PET and CT image (2 channels) (Hidden input 2: integrated detection of lymph node map by DeepMedic) Output: Map of masked segmented lesions	DSC = 0.86	Measures PET imaging features in pediatric lymphoma PET/CT scans in a fully-automated fashion. Model: an ensemble of 3 DeepMedics	GT: 1 NM physician with 11 yrs of experience segmented and determined malignancy status at lymph nodes Training/validation: 80 patients Testing: 20 patients
	Characterization (Radiomics)	Input: 3D; coronal; WB; PET-CT slices with segmented lesions Output: SUV_{max}, MTV, TLG, SA/MTV, measure of disease spread ($Dmax_{patient}$)	R = 0.95		
Weisman et al[49] (with Kieler) 2020	Detection	Input: 2D; coronal; WB; ^{18}F-FDG PET slice (from PET/CT) Output: Lymph nodes probability map contoured	TPR = 0.85 4 FP/patient	Automated detection of diseased lymph node Burden in lymphoma patients – PET/CT Model: an ensemble of 3 DeepMedics	GT: 1 NM physician with 11 yrs of experience segmented and determined malignancy status at lymph nodes Training: 58 patients Testing: 90 patients
Sibille et al,[50] 2020	Detection (localization + 4-category classification suspicious vs nonsuspicious for lung cancer or lymphoma)	Input: 2D; coronal; WB; ^{18}F-FDG PET slice fused with CT, MIP, anatomic atlas Output: Map of detected lesions classified under [suspicious or nonsuspicious] × [lung cancer or lymphoma]	For Localization: Sens = 0.81, Spec = 0.97, accuracy = 0.96 (for body parts), 0.87 (for region), 0.81 (for subregion) For Classification: FP = 1.47 (96 of 65),	^{18}F-FDG Uptake Classification in Lymphoma and Lung Cancer – using CT, PET, MIP, and atlas information	GT: 2 NM physicians annotated and segmented foci with increased ^{18}F-FDG uptake specified the anatomic location and classified. Training: 380 patients Validation:126 patients Testing: 123 patients

Author (Year)	Tasks Performed	Task Specific Input/Out	FoM	Details (Model-Related)	Details (Ground Truth, Sample Size)
			FN = 1.76 (of 65), AUC = 0.98		
Li et al,[51] 2019	End-to-end Segmentation	Input 1: 2D; axial; WB; ^{18}F-FDG PET and CT slices (6 channels) (Hidden input: single pixel probability map of lesions) Output: segmentation map of lymphoma	DSC = 0.73, Precision = 0.70, Recall = 0.81	End-to-end lymphoma segmentation – WB PET/CT Model details: DenseX-Net	GT: 3 clinicians delineated images, then verified and revised by 1 nuclear medicine expert Validation: 5-fold cross-validation Testing: 80 patients
Sadik,[52] 2019	Segmentation	Input 1: 2D axial/sagittal, coronal CT (3 channels) Input 2: manually detected liver and aorta Output: Segmentation of liver and aorta	DSC = 0.95	Automated calculation of liver and aortic ^{18}F-FDG uptake levels to serve as a reference for therapy response classification in HL and NHL CT segmentation maps were resampled to fit the ^{18}F-FDG-PET image in order to calculate SUVmedian Model details: CNN adopted from Goodfellow et al,[53] 2016	GT: 2 radiologists segmented images Training: 80 patients Validation: 6 patients
Bi et al,[24] 2017	Detection	Input: 3D; WB; coronal; ^{18}F-FDG PET with CT slices (2 channels) Output: 3D; WB; coronal; map of sFEPU regions (ie, Left, and right kidneys, bladder, brain, heart)	DSC: 0.92	Automatic detection of superpixel regions of FDG uptake of lymphoma regions Model details: MSE + CFSC	GT: 1 experienced operator manually identified ROI using PERCIST thresholding and the diagnostic report of PET-CT scan Training: 1.5 million nonmedical images, validated: 50,000 nonmedical images Testing: 11 patients
Classic Machine Learning Models					
Annunziata et al,[22] 2021	Statistical Classification (Prediction/ Prognosis)	Input: Deauville Score, qPET, MTV_0, slope (slope of a linear function of MTV) features from 3D; axial/coronal end-of-treatment than beginning-of-treatment ^{18}F-FDG PET and CT slices Output: Patient-level 2-class prediction (relapse vs progression)	PPV = 0.55, NPV = 0.83 (for DS 4–5) PPV = 0.89, NPV = 0.82 (for positive qPET) R = 0.63 (for ANN)	Assess the prognostic capacity of post-treatment ^{18}F-FDG-PET/CT in DLBCL patients Model details: multi-regression model, ANN	GT: 2 NM physicians independently evaluated using a dedicated fusion and display software Training: 26 patients Testing: 11 patients K-fold cross-validation
Lippi et al,[54] 2020	Classification (Radiophenomic)	Input: 3D; WB; coronal ^{18}F-FDG PET slices Output: Patient-level 4-class classification of malignant lymphoma (DLBCL, HL, follicular and mantle cell lymphoma)	Sens = 0.97, PPV = 0.94	Texture analysis and classification of malignant lymphoma Model details: SVM + RF	GT: 1 NM physician with 5 y of experience extracted VOIs using a 40%-threshold of SUV_{max} Evaluation: leave-one-out procedure, whereby each patient was used, in turn, as the test set, and all the other patients constituted the training set.
Mayerhoefer et al,[55] 2019	Statistical Classification (Prediction/ Prognosis)	Input: TMTVs, SUV_{mean}, TLG, entropy, and 15 other textural radiomic features Output: Patient-level 3-category metabolic risk	AUC = 0.72	Radiomic features for prediction of outcome in mantle cell lymphoma International prognostic indices for	GT: TMTV protocol used to semiautomatically construct with 41% SUV_{max} threshold Training: 75 patients Testing: 32 patients

Author (Year)	Tasks Performed	Task Specific Input/Out	FoM	Details (Model-Related)	Details (Ground Truth, Sample Size)
		(low, intermediate, high) of progression		MCL = MIPI and MIPI-b Model details: Multilayer perceptron feed-forward ANN	
Hu et al,[56] 2019	Detection	Input: 3D; WB; coronal/axial/sagittal ^{18}F-FDG PET slices and CT slices (2 Channels) Output: 3D probability map of the segmented lesion (normal organ and tumors)	Sens = 0.80, DSC = 0.59	Physical spatial characteristics of the lesions along with prior knowledge were used to optimize the technique. Density-based spatial clustering of applications with noise (DBSCAN)	GT: Segmentation ground truth obtained by 41% SUV max thresholding, no information on the physicians Testing: 48 patients
	Segmentation	Input1: 3D; WB; coronal/axial/sagittal ^{18}F-FDG PET slices and CT slices (2 Channels) Input2: Detection results Output: 3D, coronal/axial/sagittal slice with segmented normal organ and tumor lesions.	D_{ref} (DSC) = 0.74, D_{global} = 0.50, $Volume_{sup}$ = .39		
Yu et al,[57] 2018	Detection	Input: 2D; WB; coronal/axial/sagittal ^{18}F-FDG PET slices and CT slices (2 Channels) Output: probability map of detected lymphoma lesions	Sens = 1.0	Semiautomatic lymphoma detection and segmentation	GT: 1 physician contoured images Training/validation: 11 patients
	Segmentation	Input1: PET/CT images with physiologic hypermetabolic organs removed. Input2: Detection results Output: Border mask segmentation visualized on software on axial, sagittal, and coronal	DSC = 0.84	Model details: FC-CRF	
Grossiord et al,[58] 2017	Classification (Radiophenomics)	Input: 2D; coronal; ^{18}F-FDG PET and CT slices (2 channel) Output: slice-level 3-class classification (Organ, tumor, nonrelevant)	Sens = 0.65, Spec = 0.92, Accuracy = 0.86	Automated 3D lymphoma lesion segmentation - PET/CT Model details: PET/CT feature extraction, random forest classification, mixed spatial-spectral space of component-trees	GT: 1 expert manually expert segmentation at 41% SUV_{max} Training/validation : 43 patients Leave-one patient-out cross-validation for classification task
	Segmentation	Input 1: 2D; WB; coronal ^{18}F-FDG PET and CT slices (2 channel) Input 2: single cluster representative of each lesion Output: 2D, WB, coronal segmentation map	DSC = 0.75		
Desbordes et al,[59] 2016	End-to-end Segmentation	Input: 2D; WB; coronal/axial/sagittal PET slices and CT slices (2 Channels) (Hidden input: single pixel representative each lesion (based on automatic seed definition)) Output: 2D, WB, lesion segmentation map	DSC = 0.80	Cellular automata define tumor seed within ROI to obtain final segmentation by iterative growth. Model details: auto-initialization cellular automata	GT: 1 NM physician manually selected and segmented the ROI Testing: 12 patients
Lartizien et al,[60] 2014	Classification (Radiophenomics)	Input: 3D; supraclavicular; axial	AUC = 0.91	SVM classifier applied on 12 most	Evaluation set: 156 lymphomatous and 32

Author (Year)	Tasks Performed	Task Specific Input/Out	FoM	Details (Model-Related)	Details (Ground Truth, Sample Size)
		^{18}F-FDG PET slices and CT slice images Output: 2-class classification (benign vs cancer)		discriminant 1st and 2nd order textural features derived from the registered PET and contrast CT images	suspicious 25 (11 males and 14 females) baselines with B-cell lymphoma or HL.

Abbreviations: ^{18}F-FDG, ^{18}F-fluorodeoxyglucose; ANN, artificial neural network; AUC, area under curve; CFSC, class-driven feature-selection & classification; DLBCL, diffuse large B-cell lymphoma; DSC, dice similarity coefficient; ENKTL, extranodal NK/T cell lymphoma; nasal type; FC-CRF, fully connected conditional random fields; FCN, fully convolutional network; FoM, figures of merit; FROC, free-response receiver operating characteristic; ICC, intraclass correlation coefficient, JSC, Jaccard similarity coefficient; MH, metabolic heterogeneity; MHD, modified Hausdorff distance; MIP, maximum intensity projection, MLP, multilayer perception; MM, multi-regression model, MSE, multi-scale super pixel-based encoding; MTV, metabolic tumor volume; NHL, non-Hodgkin lymphoma; NM, nuclear medicine; NPV, negative predictive value; OS, overall survival; PA, percentage agreement; PERCIST, PET response criteria in solid tumors; PFS, progression-free survival; PPV, positive predictive value, PSI, prediction similarity index; R, Pearson correlation coefficient; RF, random forest; ROI, region of interest; Sens, sensitivity, sFEPU, sites of FDG excretion and physiologic FDG uptake; Spec, Specificity; SVM, Support Vector Machine, TLG, Tumor lesion glycolysis; TMTV, total Metabolic Tumor volume.

Table 2.

The mathematical definitions for the evaluation metrics used in the reviewed articles

Evaluation Measure	Mathematical Definition
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
PPV or Precision	$\frac{TP}{TP+FP}$
Dice similarity coefficient (DSC) (Synonyms: Dice similarity index, Sorensen–Dice coefficient, F1-Score, Sorensen–Dice index, Dice's Coefficient)	$\frac{2TP}{2TP+FP+FN}$
Dice $_{Ref}$ (equivalent to Dice) F(M) is defined as the number of elements in set M. G = ground truth regions composed of g_i voxels. B = set of detected lymphoma regions consisting of b_i voxels.[56]	$\frac{2F(B\cap G)}{F(G)+F(B)}$
Dice $_{Global}$[56] Same as the Dice$_{Ref}$ reference but includes false-positive regions. S is the set of all detected false regions.	$\frac{2F(B\cap G)}{F(G)+F(B)+F(S)}$
Volume $_{Sup}$[56] For evaluating the volume of the false-positive region.	$\frac{F(S)}{F(B\cup S)}$
Jaccard similarity coefficient (JSC)	$\frac{TP}{TP+FP+FN}$

Table 3.

Survey of classification, prognostication, and prediction methods based on ^{18}F-FDG-PET radiomics

Authors (Year)	Lymphoma Subtypes	Aim of Study		Radiomic Feature Information			Discriminator Used	Figures of Merit
		Input	Goal	Extraction Method	Features Used	Notable Radiomic Features		
Studies in which radiomic features were utilized for the classification of lymphoma								
Ou et al,[74] 2020	Breast lymphoma	Segmented breast tumor VOIs on ^{18}F-FDG-PET/CT	Differentiation (breast lymphoma vs breast carcinoma)	LifeX	First and second-order radiomic features	PETa and CTa models demonstrated great potential to differentiate in training and validation group	LDA	Not given for testing datasets
Xu et al,[75] 2019	Hepatic lymphoma	Segmented hepatic tumor VOIs on ^{18}F-FDG-PET	Differentiation (hepatic lymphoma vs HCC)	LifeX	6 image-based parameters and 39 texture features	Combination model of texture and image features had greater diagnostic capability	ROC analysis	AUC = 0.898
Ou et al,[76] 2019	Breast lymphoma	Segmented breast tumor VOIs on ^{18}F-FDG-PET/CT	Differentiation (breast lymphoma vs breast carcinoma)	LifeX	First and second-order radiomic features	Combination model of PET and CT features had greater diagnostic capability	Binary logistic regression	PET: AUC = 0.751. CT: AUC = 0.729; PET + CT: AUC = 0.771
Aide et al,[77] 2018	DLBCL	Axial skeleton segmented on ^{18}F-FDG-PET	Identify bone marrow involvement in DLBCL based on radiomic features from ^{18}F-FDG-PET	LifeX	4 first-order, 6 second-order and 11 third-order texture features	SkewnessH was most predicitive of lymphoma	Linear regression, ROC analysis	AUC = 0.820
Lartizien et al,[60] 2014	All types	Segmented suspicious regions of interest on^{18}-F-FDG-PET/CT	Lymphoma vs HiNA	Not reported	105 features (GLDM, GLCM, GLISZ, GLRLM, and first order)	Combination model of PET and CT features had greater diagnostic capability	SVM	Combination of CT and PET: AUC = 0.910
Studies in which radiomic features were used for the prognosis/prediction of lymphoma								
Rodriguez Taroco et al,[78] 2021	HL	Segmented tumor VOIs on ^{18}F-FDG-PET	Prediction of PFS from 18F-FDG-PET radiomic features in HL and DLBCL	Not specified	8 first-order features, 23 features from GLCM, 11 features from GLRLM, 5 features from NGLM, 3 features from the neighborhood grey-tone difference	PFS in patients with Deauville scores of 1, 2, 3, and X at initial PET was higher than that in patients with a Deauville score of 4	Univariate and multivariate Cox regression analysis	Average PFS, for patients with Deauville 4 score, of 1120 d (95% CI, 229–672)
Eertink et al,[79] 2021	DLBCL	Segmented tumor VOIs on ^{18}F-FDG-PET	Prediction of treatment outcome with first-line treatment of DLBCL from baseline ^{18}F-FDG-PET radiomic features	RaCat	Large number of morphologic and texture features were extracted	Five models were created based on radiomic features as well as clinical predictors; combination of clinical and radiomics	ROC analysis	Combined model: HR = 4.6 (95% CI, 2.6–7.9)

Authors (Year)	Lymphoma Subtypes	Aim of Study		Radiomic Feature Information			Discriminator Used	Figures of Merit
		Input	Goal	Extraction Method	Features Used	Notable Radiomic Features		
						predictors was best		
Wang et al,[80] 2020	ENKTL	Segmented tumor VOIs on ^{18}F-FDG-PET	Identify a ^{18}F-FDG-PET radiomics-based model for predicting PFS and OS in ENKTL	LifeX	41 features	Radiomics and metabolism-based models were combined to predict both PFS and OS	Univariate and multivariate Cox regression analysis	PFS: 0.788 (95% CI = 0.682–0.895) and 0.473 (P = .803) OS: 0.637 (95% CI = 0.488–0.786) and 0.730 (95% CI = 0.548–0.912)
Sun et al,[81] 2020	Primary gastric DLBCL	Segmented tumor VOIs on ^{18}F-FDG-PET	Texture analysis of 18F-PET-CT scans to predict interim response after 3–4 rounds of chemotherapy in primary gastric DLBCL	In-house software	First and second-order features	Combination of SUV_{max}, volume, and entropy in one model best predicted treatment response	Mann-Whitney U	AUC = 0.915
Aide et al,[82] 2020	DLBCL	Segmented tumor VOIs on ^{18}F-FDG-PET	Prognosticate DLBCL treated with first-line immunotherapy using radiomic features from baseline 18F-FDG-PET	LifeX	19 features	18F-FDG-PET heterogeneity of the largest lymphoma lesion is associated with 2y-event free survival (EFS)	Univariate and multivariate Cox regression analysis	EFS: HR = 7.47 (95% CI = 0.83–66.99)
Wu et al,[83] 2019	DLBCL	18F-FDG-PET/CT pre and posttreatment	Radiomics-based treatment outcome prediction model	MATLAB	GLCM, GLRLM, GLSZM	Belief-function theory-based outcome prediction outperformed than other studies	EK-NN and SVM	Therapy response: NS
Tatsumi et al,[84] 2019	FL	Segmented tumor VOIs on ^{18}F-FDG-PET	Predict response and recurrence after therapy in FL	PETSTAT	6 texture features	low gray-level zone emphasis (LGZE) in texture features predicted complete response	Logistic regression	Therapy response: AUC = 0.720; PFS: NS
Lue et al,[85] 2019	HL	Segmented tumor VOIs on ^{18}F-FDG-PET	^{18}F-FDG-PET was analyzed using radiomics to predict/prognose HL	OsiriX, CGITA, MATLAB	11 first-order, 39 higher-order, 400 wavelet features	Ann Arbor stage, GLRLM and SUV kurtosis were associated with PFS	Univariate and multivariate Cox regression analysis	PFS: HR = 6.640 (95% CI, 1.261–34.96; P = .026); OS: HR = 14.54 (95% CI, 1.808–117.0; P = .012)
Lue et al,[86] 2019	HL	Segmented tumor VOIs on ^{18}F-FDG-PET	Radiomic intratumor heterogeneity in 18F-FDG-PET to predict treatment response and survival	OsiriX, CGITA, MATLAB	7 SUV and HU, 78 second- and higher-order, 624 wavelet features	Treatment response was associated with high-intensity run emphasis (HIR) was performed on PET images and	Cox proportional hazards model, ROC curve, logistic regression	PET: Therapy response: OR = 36.4 (95% CI, 2.060–642.0, P = .014);

Authors (Year)	Lymphoma Subtypes	Aim of Study		Radiomic Feature Information			Discriminator Used	Figures of Merit
		Input	Goal	Extraction Method	Features Used	Notable Radiomic Features		
			outcomes in patients with HL			run-length nonuniformity (RLNU) of CT extracted from gray-level run-length matrix (GLRM) in high-frequency wavelets PFS was independently associate with intensity nonuniformity (INU) of PET and wavelet short-run emphasis (SRE) of CT from GLRM and Ann Arbor stage. OS was associated with zone-size nonuniformity (ZSNU) of PET from gray-level size zone matrix (GLSZM)		PFS: HR = 9.286 (95% CI, 1.341–66.28; *P* = .023); OS: HR = 41.02 (95% CI, 4.206–400.1; *P* = .001) CT:Therapy response: OR = 30.4 (95% CI, 1.700–545.0; *P* = .014); PFS: HR = 18.480 (95% CI, 1.918–178.1; *P* = .012); OS: NS
Zhou et al,[87] 2019	Primary gastric DLBCL	Segmented tumor VOIs on ^{18}F-FDG-PET	Prediction of OS and PFS from ^{18}F-FDG-PET radiomic features in primary gastric DLBCL	LifeX	44 texture features	Kurtosis, TMTV, GLNU, and HGZE were identified as independent prognostic factors	Univariate and multivariate Cox regression analysis	PET: PFS: HR = 14.642 (95% CI, 2.661–80.549; *P* = .002); OS: HR = 28.685 (95% CI, 2.067–398.152; *P* = .012) CT: PFS: HR = 11.504 (95% CI, 1.921–68.888; *P* = .007); OS: HR = 11.791 (95% CI, 1.583–87.808; *P* = .016)
Milgrom et al,[88] 2019	Mediastinal HL	Segmented nodal disease on ^{18}F-FDG-PET/CT	Predict response to therapy in mediastinal HL	MIM, IBEX	GLCM, intensity histogram, shape	A combination model of 5 most predictive features accomplished the highest AUC (SUV_{max}, TMTV, inverse variance, and 2 measures of tumor heterogeneity)	ROC analysis	AUC = 0.952
Wang et al,[89] 2019	Renal/ adrenal lymphoma	Segmented tumor VOIs on ^{18}F-FDG-PET	Prognose patients with primary renal lymphoma and	LifeX	37 texture features	GLRLM_RLNU (gray-level co-occurrence matrix run-	Univariate and multivariate Cox	OS: HR = 9.016 (95% CI, 1.041–

Authors (Year)	Lymphoma Subtypes	Aim of Study		Radiomic Feature Information			Discriminator Used	Figures of Merit
		Input	Goal	Extraction Method	Features Used	Notable Radiomic Features		
			primary adrenal lymphoma using texture features			length nonuniformity) was most predictive of OS.	regression analysis	78.112; P = .046)
Parvez et al,[90] 2018	NHL	TMTV using thresholding and radiomic features	Predict response to therapy and outcome in NHL using radiomic features extracted from 18F-FDG-PET/CT	LifeX	GLCM, NGLDM, GLRLM, GLZLM, indices from sphericity and histogram	GLNU correlated to DFS, and kurtosis correlated with OS	Univariate Cox regression analysis	Therapy response: NS; DFS: P = .013; OS: P = .035
Aide et al,[77] 2018	DLBCL	Axial skeleton segmented on ^{18}F-FDG-PET	Determine prognostic value of skeletal textural features in DLBCL	LifeX	4 first-order, 6 second-order and 11 third-order texture features	The only independent predictor of PFS was SkewnessH	ROC analysis	PFS: HR = 3.17 (95% CI, 1.00–10.04; P = .032)
Ben Bouallègue et al,[91] 2017	Bulky HL and NHL	Segmented tumor VOIs on ^{18}F-FDG-PET	Predict response to therapy in bulky HL and NHL	In-house software	Shape, texture features	SVM accounting for both texture and shape features achieved the highest ROC AUC	ROC analysis	AUC = 0.820

Abbreviations: ^{18}F-FDG-PET, ^{18}F-fluorodeoxyglucose-positron emission tomography; ACC, accuracy; AUC, area under the curve, DFS, disease-free survival; DLBCL, diffuse large B cell lymphoma; EFS, event-free survival; EK-NN, evidential k-NN; FL, follicular lymphoma; GLCM, grey-level co-occurrence matrix; GLRLM, grey-level run-length matrix; GLSZM, grey-level size-zone matrix; GLZLM, grey-level zone length matrix; HL, Hodgkin's lymphomas; HR, hazard ratio; LDA, linear discriminant analysis; MCL, mantle cell lymphoma; NGLDM, neighborhood grey-level different matrix; NHL, non-Hodgkin's lymphomas; NS, not significant; OR, odds ratio; OS, overall survival; PFS, progression-free survival; ROC, receiver operating characteristic; RUN, run-length matrix; SEN, sensitivity; SPE, specificity; SVM, support vector machine; TF, texture features; VOI, volume of interest.