Methodological evaluation of original articles on radiomics and machine learning for outcome prediction based on positron emission tomography (PET)

Methodische Bewertung von Originalartikeln zu Radiomics und Machine Learning für Outcome-Vorhersagen basierend auf der Positronen-Emissions-Tomografie (PET)

Authors

Julian Manuel Michael Rogasch[1, 2], Kuangyu Shi[3], David Kersting[4], Robert Seifert[4]

Affiliations

1 Department of Nuclear Medicine, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
2 Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin
3 Department of Nuclear Medicine, Inselspital University Hospital Bern, Bern, Switzerland
4 Department of Nuclear Medicine, University Hospital Essen, Essen, Germany

Key words

radiomics, positron emission tomography, artificial intelligence, machine learning, TRIPOD, outcome prediction

received 15.9.2023
accepted 25.10.2023

Bibliography

Nuklearmedizin 2023; 62: 361–369
DOI 10.1055/a-2198-0545
ISSN 0029-5566

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence

M.D. Julian Manuel Michael Rogasch
Department of Nuclear Medicine, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Augustenburger Platz 1, 13353 Berlin, Germany
julian.rogasch@charite.de

Additional material is available at https://doi.org/10.1055/a-2198-0545.

ABSTRACT

Aim Despite a vast number of articles on radiomics and machine learning in positron emission tomography (PET) imaging, clinical applicability remains limited, partly owing to poor methodological quality. We therefore systematically investigated the methodology described in publications on radiomics and machine learning for PET-based outcome prediction.

Methods A systematic search for original articles was run on PubMed. All articles were rated according to 17 criteria proposed by the authors. Criteria with > 2 rating categories were binarized into "adequate" or "inadequate". The association between the number of "adequate" criteria per article and the date of publication was examined.

Results One hundred articles were identified (published between 07/2017 and 09/2023). The median proportion of articles per criterion that were rated "adequate" was 65 % (range: 23–98 %). Nineteen articles (19 %) mentioned neither a test cohort nor cross-validation to separate training from testing. The median number of criteria with an "adequate" rating per article was 12.5 out of 17 (range, 4–17), and this did not increase with later dates of publication (Spearman's rho, 0.094; p = 0.35). In 22 articles (22 %), less than half of the items were rated "adequate". Only 8 % of articles published the source code, and 10 % made the dataset openly available.

Conclusion Among the articles investigated, methodological weaknesses have been identified, and the degree of compliance with recommendations on methodological quality and reporting shows potential for improvement. Better adherence to established guidelines could increase the clinical significance of radiomics and machine learning for PET-based outcome prediction and finally lead to the widespread use in routine clinical practice.

Introduction

In addition to its clinical value for tumor detection and staging, metabolic and/or molecular information derived from positron emission tomography (PET) imaging can facilitate the prognostication of survival and the prediction of treatment outcomes in various tumor types [1, 2, 3]. Traditionally employed image-derived features comprise standardized uptake values (SUV) and metabolic tumor volume (MTV) as well as composite metrics like total lesion glycolysis. These metrics are most commonly derived from manually or semi-automatically delineated regions of interest. In recent years, radiomics and machine learning-based prediction models have been increasingly employed to enhance the prognostic or predictive value of PET imaging by leveraging textural information and patterns that are not directly accessible to human readers [4, 5]. However, the increasing complexity and feature number of such approaches, compared to the sparsity of manually derived image features, brings with it a higher risk of obtaining results that are either biased or not reproducible. Finally, despite the variety of prognostic radiomics and machine learning-based models published so far, broad clinical applicability has still not been achieved.

Different guidelines and recommendations have been published to define standards for radiomics and multivariable prediction approaches [6, 7, 8]. A common characteristic of these guidelines is that they propose standard practices related to both methodological quality and the transparency of reporting. This is a necessity, as readers, journal editors and reviewers of radiomics and machine learning articles should be enabled to fully assess the methodological approach. Ideally, a full description of methodological details should enable readers to repeat the experiment themselves. For example, the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) statement has set standards for categorizing studies on multivariable models based on the type of internal or external validation that was used. The European Association of Nuclear Medicine (EANM) and Society of Nuclear Medicine and Molecular Imaging (SNMMI) have recently published joint recommendations on radiomics research in nuclear medicine [6]. Lambin *et al.* have proposed a radiomics quality score (RQS) to assess the methodological quality of radiomics studies [7].

Using the TRIPOD criteria and RQS, Park *et al.* rated 77 radiomics studies in oncology and found that at 9.4, the mean RQS was only 26 % of the maximum score of 36. Furthermore, TRIPOD recommendations were followed to a varying degree (2.6 % to 100 %). However, only 4 % of the evaluated articles investigated PET, and only 20 % of the studies focused on prognostic tasks while the majority assessed diagnostic applications [9].

The RQS defines several aspects that should be considered in radiomics analyses but does not rate or quantify the appropriateness of the implementation of methods. TRIPOD recommendations are primarily focused on the transparency of reporting rather than provision of methodological guidance and are not specifically optimized for radiomics analyses.

On this basis, we set out to further explore the methodological approaches and reporting transparency of radiomics and machine learning articles that focus on PET-based outcome prediction using our own set of objective and subjective rating criteria. Our assessment was designed to resemble that of independent reviewers, journal editors or readers who would critically judge not only the concordance with specific methodological aspects but also the overall appropriateness of methods in the individual study context and the overall reliability of results and conclusions. Furthermore, we examined whether ratings improved with later dates of publication.

▶ **Table 1** Search prompts and total number of article results.

Search prompt	Number of articles
(artificial intelligence[Title/Abstract]) AND (PET[Title/Abstract]) AND (outcome[Title/Abstract]) AND (prediction[Title/Abstract])	12
(neural network[Title/Abstract]) AND (PET[Title/Abstract]) AND (outcome[Title/Abstract]) AND (prediction[Title/Abstract])	10
(automated[Title/Abstract]) AND (PET[Title/Abstract]) AND (outcome[Title/Abstract]) AND (prediction[Title/Abstract])	13
(neural network[Title/Abstract]) AND (PET[Title/Abstract]) AND (response[Title/Abstract]) AND (prediction[Title/Abstract])	8
neural network PSMA prediction	6
deep learning PSMA prediction	6
(neural network[Title/Abstract]) AND (PET[Title/Abstract]) AND (survival[Title/Abstract])	36
(artificial intelligence[Title/Abstract]) AND (PET[Title/Abstract]) AND (survival[Title/Abstract])	38
(radiomics[Title/Abstract]) AND (PET[Title/Abstract]) AND (outcome[Title/Abstract])	155

Material and Methods

Search strategy

Original articles were identified by a single person (JMMR) through a search on PubMed on July 15 and 16, 2023. Search prompts and results are given in ▶ **Table 1**. Only original articles in English that investigated outcome prediction with PET using radiomics / textural features and/or machine learning methods (including neural networks) for classification/regression and/or to extract image features were considered for this analysis. Articles that investigated only a classification task (e. g., "predicting" histological subtypes) were not considered eligible. Date of publication was not restricted.

Rating criteria and rating process

Rating criteria and their categories / scales were created by JMMR and RS in consensus based on the RQS [7] and TRIPOD recommendations [8]. A final set of seventeen criteria was drawn up (▶ **Table 2**). Some criteria were designed as binary items (e. g.,

▶ **Table 2** List of rating criteria, their categories and the percentage of articles rated as "adequate" after binarization of the rating categories. Details on the precise definitions of all categories can be found in the **Supplemental Material**.

Rating criteria	Categories				Adequate articles (%)
	Adequate		Inadequate		
Validity of the results					
Homogeneous patient cohort?	Yes	Rather yes	Rather no	No	64
Presence of selection bias?	No	Rather no	Rather yes	Yes	88
Is the frequency of classes balanced?	Yes	Rather yes	Rather no	No	65
Training and testing split consistently (to prevent data leakage)?	Yes	Rather yes	Rather no	No	64
Clearly defined research question(s)?	Yes	Rather yes	Rather no	No	98
Clear main/primary endpoint for model training?	Yes	Rather yes	Rather no	No	86
Adequate statistical method for the primary endpoint?	Yes	Rather yes	Rather no	No	77
Comparison with established biomarkers?	Yes		No		65
Machine learning description informative?	Yes	Rather yes	Rather no	No	64
Generalizability of the results					
Risk of overfitting?	Low	Average	High		62
Separate test cohort / cross-validation present?	Yes		No		81
Robustness of results reported (e. g., resampling or confidence interval)?	Yes		No		39
Independent/external cohort used for testing?	Yes		No		23
Results and conclusion					
Is the presentation of results comprehensible?	Yes	Rather yes	Rather no	No	90
Are the results informative?	Yes	Rather yes	Rather no	No	72
Adequate conclusion with regard to validity?	Yes	Rather yes	Rather no	No	73
Adequate conclusion with regard to generalizability?	Yes	Rather yes	Rather no	No	59

whether a separate test cohort was analyzed). Other, more subjective items were rated on a 3- or 4-point scale. To facilitate the interpretation of the results, all rating criteria were binarized into "adequate" or "inadequate" rating categories (▶ **Table 2**). A description of how individual categories for each item were defined can be found in the **Supplemental Material**.

A single rater (JMMR) assessed all original articles based on the seventeen criteria. If a certain criterion could not be rated (because the article did not contain the required information), it was rated as "unclear" and was also regarded as "inadequate". Furthermore, the sample size of the training cohort/folds and test cohort/folds was noted and whether data collection included prospective or multicentric data. Whether the dataset and/or the machine learning code had been published on a public repository was also evaluated. The TRIPOD study type was identified, based on the methodology applied in the articles. On average, around 20 minutes were required to rate an article.

Statistical analysis

Based on the Shapiro-Wilk test, a non-normal data distribution was assumed, and the median, interquartile range (IQR), and range were presented. The association between the date of publication and number of criteria with an "adequate" rating per article was analyzed by Spearman correlation using SPSS version 29.0.0.0 (IBM Corporation, Armond, NY, USA). Statistical significance was assumed at $\alpha = 0.05$.

Results

Description of the original articles

A total of 107 original articles were identified through the search process. Three of these were excluded as no full text version could be accessed. One publication was excluded because it investigated only one single PET feature, while another article was not included because the declared aim was solely to validate a previously published model. Two publications examining PET in the context of Alzheimer's disease or amyotrophic lateral sclerosis were excluded, as these differed fundamentally from all the other articles where the focus was on oncology.

The remaining articles (n = 100), published between July 2017 and September 2023, were included in the subsequent analysis (▶ **Table 3**). The cancer entities investigated comprised head and neck cancer (n = 26) [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,

▶ **Table 3** General characteristics of the articles analyzed.

Characteristics	Number of articles
Total	100
Type of image analysis (categories based on [6])	
Only hand-crafted radiomics	50
Radiomics and machine learning for prediction ("hybrid radiomics")	46
Only machine learning for prediction	4
Year of publication	
2017	1
2018	7
2019	13
2020	15
2021	20
2022	29
2023	15
Prospective data included	19
Multicentric data included	30
TRIPOD study type	
1a	15
1b	26
2a	31
2b	8
3	16
Unclear (considered equivalent to type 1a)	4

22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35], lung cancer (n = 20) [36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55], gastrointestinal or hepatobiliary tumors (n = 15) [56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70], lymphoma (n = 13) [3, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82], gynecological tumors (n = 7) [83, 84, 85, 86, 87, 88, 89], prostate cancer (n = 5) [90, 91, 92, 93, 94], breast cancer (n = 3) [95, 96, 97], brain tumors (n = 3) [98, 99, 100], sarcoma (n = 2) [101, 102], melanoma (n = 2) [103, 104], multiple myeloma (n = 2) [105, 106], salivary gland tumors (n = 1) [107] and pleural mesothelioma (n = 1) [108], respectively.

The median sample size of the training cohort/folds was 86 (IQR, 49–172; range, 20–1049). Nineteen articles (19 %) reported neither a test cohort nor cross-validation to separate training from testing data. In the other 81 articles, the median sample size of the test cohort (fold in case of cross-validation) was 40 (IQR, 20–85; range, 1–887). A test sample size of one in two articles was due to bootstrapping. This is considered adequate by the TRIPOD recommendations (study type 1b).

In 8 articles (8 %), the machine learning code was published, while in 10 (10 %) the datasets were made publically available.

Rating results

The median proportion of articles per criterion with "adequate" rating was 65 % (IQR, 63–84), ranging from 23 % (criterion: "Independent/external cohort used for testing?") to 98 % (criterion: "Clearly defined research question(s)?").

The median number of criteria with an "adequate" rating per article was 12.5 (IQR, 9–14; range, 4–17). Two articles [12, 42] were rated "adequate" in all 17 criteria, whereas 22 articles (22 %) had an "adequate" rating in less than half of the items (≤ 8 of 17 items). The number of criteria with an "adequate" rating per article was not associated with the date of publication (Spearman's rho, 0.094; p = 0.35; ▶ **Fig. 1**).

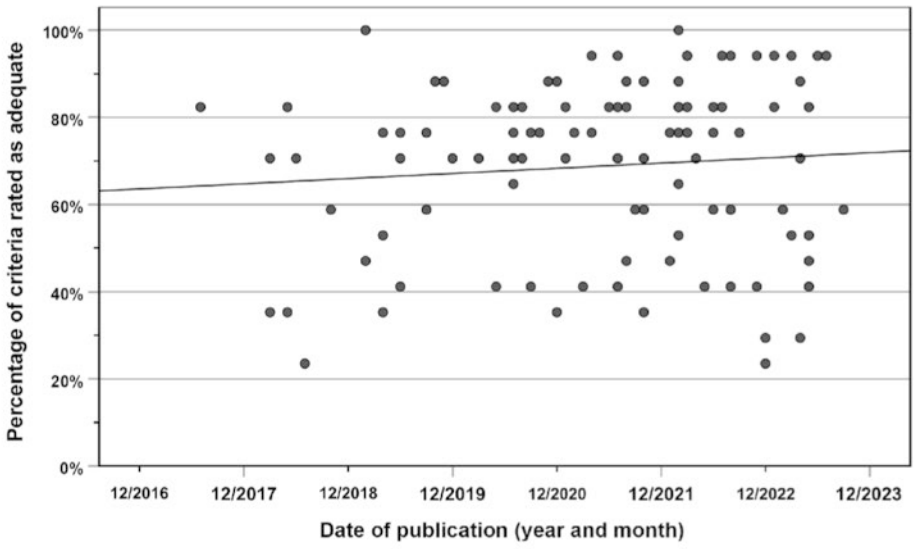
Detailed results item-by-item for all articles are available as open data in the Zenodo repository https://doi.org/10.5281/zenodo.8284734.

Discussion

The aim of this investigation was to analyze the appropriateness of methodological approaches and their description in studies on radiomics or machine learning analysis in PET-based outcome prediction. To this end, a multi-dimensional set of seventeen items was drawn up and applied to 100 articles retrieved from a systematic PubMed search. On average, we found that per criterion a median of 65 % of articles achieved an "adequate" rating.

The fraction of 65 % adequate articles per criterion based on the comprehensive review may seem high, but it should be noted that the criteria and rating categories used were designed to reflect scientific reality and are not an idealized situation. This is underlined by the observation that two articles [12, 42] achieved an "adequate" rating for all 17 criteria assessed by us. Like us, Park *et al.* examined 77 articles on radiomics but reported that at 9.4, the average RQS was only 26.1 % of the maximum of 36 [9]. As we used different rating criteria, these results can only be compared at the level of individual items. Like that of Park *et al.*, our analysis showed that most of the studies reserved a separate cohort or used cross-validation for testing (here: 81 %; Park *et al.*: 90 %). However, it is worth noting that 19 % did not do so, which means that the corresponding results are purely exploratory and cannot be generalized. Neither should they be applied in clinical routine practice without prior validation. Most articles used adequate statistical (discrimination) methods for the primary endpoint, such as area under the curve (AUC) or concordance index (77 %). Park *et al.* reported even higher adherence (99 %), probably because most studies analyzed by Park *et al.* were diagnostic studies that usually use binary endpoints and can rely on the AUC. We also found that resampling was employed to examine the robustness of results by only 39 % of the studies examined (this was similar to Park *et al.*: 30 %). Validation/testing with an independent (external) dataset was unfortunately not frequently observed (23 %), and datasets or source codes were rarely made available (10 and 8 %, respectively), which is in line with Park *et al.* (external validation: 18 %; open science: 4 %).

The TRIPOD statement and RQS provide a list of methods or report elements that should be observed. However, beyond simply following such a checklist of recommendations, authors should



▶ **Fig. 1** Correlation between date of publication (grid: months) and percentage of criteria per article that were rated as "adequate" (the maximum of 17 criteria would correspond to 100 %). Solid line: linear fit.

also report how rigorously the methods were implemented because simply using tools such as cross-validation or feature selection does not guarantee their proper implementation to completely prevent data leakage. We think that authors should make every effort to enable readers to understand the way that such crucial methodological steps were actually realized. Positive examples included articles that used flow charts to visualize the separation of training/test and other methodological steps. However, we found that key aspects were often not sufficiently covered by the articles ("informative" machine learning description in only 64 % of articles).

One example of a RQS criterion that may – formally speaking – be fulfilled by many articles is the use of a multivariable analysis such as Cox regression. In many of the investigated articles, Cox regression was used for feature selection or to investigate the independent value of predictors. However, proper implementation requires that weights are kept unchanged between training and testing. Otherwise, retraining would invalidate the training/testing separation approach. Unfortunately, this was often not stated. In general, descriptions of machine learning methods are rarely given in sufficient detail (or supported by open data code) to allow other researchers to reliably reproduce the selected features and final results.

In studies with multiple endpoints (e. g., progression yes/no, progression-free survival, and overall survival), the authors should specify the endpoint used for training. If a binary endpoint was used, several of the investigated studies reported multiple performance metrics (e. g., AUC and accuracy) and omitted to state clearly, which of the mentioned performance metrics was used to gauge model training.

We identified aspects in methodology or reporting that are not (fully) covered by the TRIPOD statement or the RQS: the balance/imbalance of outcomes, an overall assessment of the risk of overfitting (in the light of sample size, feature selection methods, and loss of performance from training to testing) or the adequateness of conclusions. We strongly feel that these criteria are also important factors for judging the validity of radiomics results. The proposed set of criteria is therefore useful for reviewers of radiomics and machine learning articles for PET-based predictions.

In order to objectify our criteria, we have included descriptions of all subjective rating categories (**Supplemental Material**), which should enable readers to understand and interpret our findings. We also point out the limitations of our approach and state which standards were required for an "adequate" rating. We hope that these descriptions will improve the reproducibility of the rating. It is furthermore worth noting that a single person rated the articles, and subjective ratings may be influenced by personal expectations and demands on the quality of radiomics papers. Furthermore, inter-rater variability could not be determined. Our analysis and its transferability to other analyses is biased, because we used our own set of rating criteria and categories. However, our findings on specific items that are also part of the RQS were comparable to Park *et al.* [9], indicating the validity of the single observer assessment at least for these criteria.

Regardless of the potential subjectivity of the evaluation criteria, the observation remains that the average rating, which is a measure of the quality of the paper, did not change (improve) over the time period 07/2017 to 09/2023. Notably, the TRIPOD statement was published in 2015 [8] and the RQS in 2017 [7]. Still, the impact of these recommendations on how radiomics and machine learning analyses are performed and reported in the context of outcome prediction with PET appears to be limited. Considering that the clinical applicability of radiomics analyses remains low, despite the large number of published models, improvements are clearly needed.

Conclusion

The methodological quality and reporting transparency of the investigated papers on radiomics and machine learning for PET-based outcome prediction often appeared inadequate. Authors and reviewers of such articles should aim at enabling the reader to reproduce the results and to this end, certain critical quality criteria must be met (e. g., clear identification of event rates, the endpoint and primary performance metric used for model training, and a clear description of the separation of training/testing to identify potential data leakage). Better adherence to previously published recommendations is imperative to finally enable the widespread use of radiomics and machine learning in routine clinical practice.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Berghmans T, Dusart M, Paesmans M et al. Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. J Thorac Oncol 2008; 3: 6–12. doi:10.1097/JTO.0b013e31815e6d6b

[2] Seifert R, Rasul S, Seitzer K et al. A Prognostic Risk Score for Prostate Cancer Based on PSMA PET-derived Organ-specific Tumor Volumes. Radiology 2023; 307: e222010. doi:10.1148/radiol.222010

[3] Eertink JJ, Zwezerijnen GJC, Heymans MW et al. Baseline PET radiomics outperforms the IPI risk score for prediction of outcome in diffuse large B-cell lymphoma. Blood 2023; 141: 3055–3064. doi:10.1182/blood.2022018558

[4] Yousefirizi F, Pierre Decazes null, Amyar A et al. AI-Based Detection, Classification and Prediction/Prognosis in Medical Imaging:: Towards Radiophenomics. PET Clin 2022; 17: 183–212. doi:10.1016/j.cpet.2021.09.010

[5] Philip MM, Welch A, McKiddie F et al. A systematic review and meta-analysis of predictive and prognostic models for outcome prediction using positron emission tomography radiomics in head and neck squamous cell carcinoma patients. Cancer Med 2023; 12: 16181–16194. doi:10.1002/cam4.6278

[6] Hatt M, Krizsan AK, Rahmim A et al. Joint EANM/SNMMI guideline on radiomics in nuclear medicine : Jointly supported by the EANM Physics Committee and the SNMMI Physics, Instrumentation and Data Sciences Council. Eur J Nucl Med Mol Imaging 2023; 50: 352–375. doi:10.1007/s00259-022-06001-6

[7] Lambin P, Leijenaar RTH, Deist TM et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017; 14: 749–762. doi:10.1038/nrclinonc.2017.141

[8] Collins GS, Reitsma JB, Altman DG et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015; 350: g7594. doi:10.1136/bmj.g7594

[9] Park JE, Kim D, Kim HS et al. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. Eur Radiol 2020; 30: 523–536. doi:10.1007/s00330-019-06360-z

[10] Wang K, Zhou Z, Wang R et al. A multi-objective radiomics model for the prediction of locoregional recurrence in head and neck squamous cell cancer. Med Phys 2020; 47: 5392–5400. doi:10.1002/mp.14388

[11] Lv W, Xu H, Han X et al. Context-Aware Saliency Guided Radiomics: Application to Prediction of Outcome and HPV-Status from Multi-Center PET/CT Images of Head and Neck Cancer. Cancers (Basel) 2022; 14. doi:10.3390/cancers14071674

[12] Le WT, Vorontsov E, Romero FP et al. Cross-institutional outcome prediction for head and neck cancer patients using self-attention neural networks. Sci Rep 2022; 12: 3183. doi:10.1038/s41598-022-07034-5

[13] Cheng N-M, Yao J, Cai J et al. Deep Learning for Fully Automated Prediction of Overall Survival in Patients with Oropharyngeal Cancer Using FDG-PET Imaging. Clin Cancer Res 2021; 27: 3948–3959. doi:10.1158/1078-0432.CCR-20-4935

[14] Zhao X, Liang Y-J, Zhang X et al. Deep learning signatures reveal multiscale intratumor heterogeneity associated with biological functions and survival in recurrent nasopharyngeal carcinoma. Eur J Nucl Med Mol Imaging 2022; 49: 2972–2982. doi:10.1007/s00259-022-05793-x

[15] Salmanpour MR, Rezaeijo SM, Hosseinzadeh M et al. Deep versus Handcrafted Tensor Radiomics Features: Prediction of Survival in Head and Neck Cancer Using Machine Learning and Fusion Techniques. Diagnostics (Basel) 2023; 13. doi:10.3390/diagnostics13101696

[16] Wang C, Liu C, Chang Y et al. Dose-Distribution-Driven PET Image-Based Outcome Prediction (DDD-PIOP): A Deep Learning Study for Oropharyngeal Cancer IMRT Application. Front Oncol 2020; 10: 1592. doi:10.3389/fonc.2020.01592

[17] Beichel RR, Ulrich EJ, Smith BJ et al. FDG PET based prediction of response in head and neck cancer treatment: Assessment of new quantitative imaging features. PLoS One 2019; 14: e0215465. doi:10.1371/journal.pone.0215465

[18] Lv W, Zhou Z, Peng J et al. Functional-structural sub-region graph convolutional network (FSGCN): Application to the prognosis of head and neck cancer with PET/CT imaging. Comput Methods Programs Biomed 2023; 230: 107341. doi:10.1016/j.cmpb.2023.107341

[19] Han K, Joung JF, Han M et al. Locoregional Recurrence Prediction Using a Deep Neural Network of Radiological and Radiotherapy Images. J Pers Med 2022; 12. doi:10.3390/jpm12020143

[20] Starke S, Zwanenburg A, Leger K et al. Longitudinal and Multimodal Radiomics Models for Head and Neck Cancer Outcome Prediction. Cancers (Basel) 2023; 15. doi:10.3390/cancers15030673

[21] Zhong J, Frood R, Brown P et al. Machine learning-based FDG PET-CT radiomics for outcome prediction in larynx and hypopharynx squamous cell carcinoma. Clin Radiol 2021; 76: 78.e9–78.e17. doi:10.1016/j.crad.2020.08.030

[22] Lv W, Ashrafinia S, Ma J et al. Multi-Level Multi-Modality Fusion Radiomics: Application to PET and CT Imaging for Prognostication of Head and Neck Cancer. IEEE J Biomed Health Inform 2020; 24: 2268–2277. doi:10.1109/JBHI.2019.2956354

[23] Dmytriw AA, Ortega C, Anconina R et al. Nasopharyngeal Carcinoma Radiomic Evaluation with Serial PET/CT: Exploring Features Predictive of Survival in Patients with Long-Term Follow-Up. Cancers (Basel) 2022; 14. doi:10.3390/cancers14133105

[24] Zhang Q, Wang K, Zhou Z et al. Predicting local persistence/recurrence after radiation therapy for head and neck cancer from PET/CT using a multi-objective, multi-classifier radiomics model. Front Oncol 2022; 12: 955712. doi:10.3389/fonc.2022.955712

[25] Gu B, Meng M, Bi L et al. Prediction of 5-year progression-free survival in advanced nasopharyngeal carcinoma with pretreatment PET/CT using multi-modality deep learning-based radiomics. Front Oncol 2022; 12: 899351. doi:10.3389/fonc.2022.899351

[26] Folkert MR, Setton J, Apte AP et al. Predictive modeling of outcomes following definitive chemoradiotherapy for oropharyngeal cancer based

on FDG-PET image characteristics. Phys Med Biol 2017; 62: 5327–5343. doi:10.1088/1361-6560/aa73cc

[27] Martens RM, Koopman T, Noij DP et al. Predictive value of quantitative (18)F-FDG-PET radiomics analysis in patients with head and neck squamous cell carcinoma. EJNMMI Res 2020; 10: 102. doi:10.1186/s13550-020-00686-2

[28] Lv W, Yuan Q, Wang Q et al. Radiomics Analysis of PET and CT Components of PET/CT Imaging Integrated with Clinical Parameters: Application to Prognosis for Nasopharyngeal Carcinoma. Mol Imaging Biol 2019; 21: 954–964. doi:10.1007/s11307-018-01304-3

[29] Ger RB, Zhou S, Elgohari B et al. Radiomics features of the primary tumor fail to improve prediction of overall survival in large cohorts of CT- and PET-imaged head and neck cancer patients. PLoS One 2019; 14: e0222509. doi:10.1371/journal.pone.0222509

[30] Marschner SN, Lombardo E, Minibek L et al. Risk Stratification Using (18)F-FDG PET/CT and Artificial Neural Networks in Head and Neck Cancer Patients Undergoing Radiotherapy. Diagnostics (Basel) 2021; 11. doi:10.3390/diagnostics11091581

[31] Sörensen A, Carles M, Bunea H et al. Textural features of hypoxia PET predict survival in head and neck cancer during chemoradiotherapy. Eur J Nucl Med Mol Imaging 2020; 47: 1056–1064. doi:10.1007/s00259-019-04609-9

[32] Nakajo M, Kawaji K, Nagano H et al. The Usefulness of Machine Learning-Based Evaluation of Clinical and Pretreatment [(18)F]-FDG-PET/CT Radiomic Features for Predicting Prognosis in Hypopharyngeal Cancer. Mol Imaging Biol 2023; 25: 303–313. doi:10.1007/s11307-022-01757-7

[33] Nakajo M, Nagano H, Jinguji M et al. The usefulness of machine-learning-based evaluation of clinical and pretreatment (18)F-FDG-PET/CT radiomic features for predicting prognosis in patients with laryngeal cancer. Br J Radiol 2023; 96: 20220772. doi:10.1259/bjr.20220772

[34] Wang K, Dohopolski M, Zhang Q et al. Towards reliable head and neck cancers locoregional recurrence prediction using delta-radiomics and learning with rejection option. Med Phys 2023; 50: 2212–2223. doi:10.1002/mp.16132

[35] Carles M, Fechter T, Grosu AL et al. (18)F-FMISO-PET Hypoxia Monitoring for Head-and-Neck Cancer Patients: Radiomics Analyses Predict the Outcome of Chemo-Radiotherapy. Cancers (Basel) 2021; 13. doi:10.3390/cancers13143449

[36] Nie P, Yang G, Wang N et al. Additional value of metabolic parameters to PET/CT-based radiomics nomogram in predicting lymphovascular invasion and outcome in lung adenocarcinoma. Eur J Nucl Med Mol Imaging 2021; 48: 217–230. doi:10.1007/s00259-020-04747-5

[37] Mattonen SA, Davidzon GA, Benson J et al. Bone Marrow and Tumor Radiomics at (18)F-FDG PET/CT: Impact on Outcome Prediction in Non-Small Cell Lung Cancer. Radiology 2019; 293: 451–459. doi:10.1148/radiol.2019190357

[38] van Timmeren JE, Carvalho S, Leijenaar RTH et al. Challenges and caveats of a multi-center retrospective radiomics study: an example of early treatment response assessment for NSCLC patients using FDG-PET/CT radiomics. PLoS One 2019; 14: e0217536. doi:10.1371/journal.pone.0217536

[39] Sepehri S, Tankyevych O, Upadhaya T et al. Comparison and Fusion of Machine Learning Algorithms for Prospective Validation of PET/CT Radiomic Features Prognostic Value in Stage II-III Non-Small Cell Lung Cancer. Diagnostics (Basel) 2021; 11. doi:10.3390/diagnostics11040675

[40] Yang B, Zhong J, Zhong J et al. Development and Validation of a Radiomics Nomogram Based on (18)F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography and Clinicopathological Factors to Predict the Survival Outcomes of Patients With Non-Small Cell Lung Cancer. Front Oncol 2020; 10: 1042. doi:10.3389/fonc.2020.01042

[41] Tankyevych O, Trousset F, Latappy C et al. Development of Radiomic-Based Model to Predict Clinical Outcomes in Non-Small Cell Lung Cancer Patients Treated with Immunotherapy. Cancers (Basel) 2022; 14. doi:10.3390/cancers14235931

[42] Arshad MA, Thornton A, Lu H et al. Discovery of pre-therapy 2-deoxy-2-(18)F-fluoro-D-glucose positron emission tomography-based radiomics classifiers of survival outcome in non-small-cell lung cancer patients. Eur J Nucl Med Mol Imaging 2019; 46: 455–466. doi:10.1007/s00259-018-4139-4

[43] Buizza G, Toma-Dasu I, Lazzeroni M et al. Early tumor response prediction for lung cancer patients using novel longitudinal pattern features from sequential PET/CT image scans. Phys Med 2018; 54: 21–29. doi:10.1016/j.ejmp.2018.09.003

[44] Afshar P, Mohammadi A, Tyrrell PN et al. [Formula: see text]: deep learning-based radiomics for the time-to-event outcome prediction in lung cancer. Sci Rep 2020; 10: 12366. doi:10.1038/s41598-020-69106-8

[45] Cui S, Ten Haken RK, El Naqa I. Integrating Multiomics Information in Deep Learning Architectures for Joint Actuarial Outcome Prediction in Non-Small Cell Lung Cancer Patients After Radiation Therapy. Int J Radiat Oncol Biol Phys 2021; 110: 893–904. doi:10.1016/j.ijrobp.2021.01.042

[46] Forouzannezhad P, Maes D, Hippe DS et al. Multitask Learning Radiomics on Longitudinal Imaging to Predict Survival Outcomes following Risk-Adaptive Chemoradiation for Non-Small Cell Lung Cancer. Cancers (Basel) 2022; 14. doi:10.3390/cancers14051228

[47] Zuo Y, Liu Q, Li N et al. Optimal (18)F-FDG PET/CT radiomics model development for predicting EGFR mutation status and prognosis in lung adenocarcinoma: a multicentric study. Front Oncol 2023; 13: 1173355. doi:10.3389/fonc.2023.1173355

[48] Amini M, Hajianfar G, Hadadi Avval A et al. Overall Survival Prognostic Modelling of Non-small Cell Lung Cancer Patients Using Positron Emission Tomography/Computed Tomography Harmonised Radiomics Features: The Quest for the Optimal Machine Learning Algorithm. Clin Oncol (R Coll Radiol) 2022; 34: 114–127. doi:10.1016/j.clon.2021.11.014

[49] Yang L, Xu P, Li M et al. PET/CT Radiomic Features: A Potential Biomarker for EGFR Mutation Status and Survival Outcome Prediction in NSCLC Patients Treated With TKIs. Front Oncol 2022; 12: 894323. doi:10.3389/fonc.2022.894323

[50] Ahn HK, Lee H, Kim SG et al. Pre-treatment (18)F-FDG PET-based radiomics predict survival in resected non-small cell lung cancer. Clin Radiol 2019; 74: 467–473. doi:10.1016/j.crad.2019.02.008

[51] Huang B, Sollee J, Luo YH et al. Prediction of lung malignancy progression and survival with machine learning based on pre-treatment FDG-PET/CT. EBioMedicine 2022; 82: 104127. doi:10.1016/j.ebiom.2022.104127

[52] Oliveira C, Amstutz F, Vuong D et al. Preselection of robust radiomic features does not improve outcome modelling in non-small cell lung cancer based on clinical routine FDG-PET imaging. EJNMMI Res 2021; 11: 79. doi:10.1186/s13550-021-00809-3

[53] Oikonomou A, Khalvati F, Tyrrell PN et al. Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. Sci Rep 2018; 8: 4003. doi:10.1038/s41598-018-22357-y

[54] Kirienko M, Sollini M, Corbetta M et al. Radiomics and gene expression profile to characterise the disease and predict outcome in patients with lung cancer. Eur J Nucl Med Mol Imaging 2021; 48: 3643–3655. doi:10.1007/s00259-021-05371-7

[55] Ventura D, Schindler P, Masthoff M et al. Radiomics of Tumor Heterogeneity in (18)F-FDG-PET-CT for Predicting Response to Immune Checkpoint Inhibition in Therapy-Naïve Patients with Advanced Non-Small-Cell Lung Cancer. Cancers (Basel) 2023; 15. doi:10.3390/cancers15082297

[56] Anconina R, Ortega C, Metser U et al. Combined 18 F-FDG PET/CT Radiomics and Sarcopenia Score in Predicting Relapse-Free Survival and Overall Survival in Patients With Esophagogastric Cancer. Clin Nucl Med 2022; 47: 684–691. doi:10.1097/RLU.0000000000004253

[57] Chen YH, Lue KH, Chu SC et al. Combining the radiomic features and traditional parameters of (18)F-FDG PET with clinical profiles to improve prognostic stratification in patients with esophageal squamous cell car-

cinoma treated with neoadjuvant chemoradiotherapy and surgery. Ann Nucl Med 2019; 33: 657–670. doi:10.1007/s12149-019-01380-7

[58] Yang CK, Yeh JCY, Yu WH et al. Deep Convolutional Neural Network-Based Positron Emission Tomography Analysis Predicts Esophageal Cancer Outcome. J Clin Med 2019; 8. doi:10.3390/jcm8060844

[59] Foley KG, Shi Z, Whybra P et al. External validation of a prognostic model incorporating quantitative PET image features in oesophageal cancer. Radiother Oncol 2019; 133: 205–212. doi:10.1016/j.radonc.2018.10.033

[60] Lovinfosse P, Polus M, Van Daele D et al. FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. Eur J Nucl Med Mol Imaging 2018; 45: 365–375. doi:10.1007/s00259-017-3855-5

[61] Fiz F, Masci C, Costa G et al. PET/CT-based radiomics of mass-forming intrahepatic cholangiocarcinoma improves prediction of pathology data and survival. Eur J Nucl Med Mol Imaging 2022; 49: 3387–3400. doi:10.1007/s00259-022-05765-1

[62] Murakami Y, Kawahara D, Tani S et al. Predicting the Local Response of Esophageal Squamous Cell Carcinoma to Neoadjuvant Chemoradiotherapy by Radiomics with a Machine Learning Method Using (18)F-FDG PET Images. Diagnostics (Basel) 2021; 11. doi:10.3390/diagnostics11061049

[63] Beukinga RJ, Hulshoff JB, Mul VEM et al. Prediction of Response to Neoadjuvant Chemotherapy and Radiation Therapy with Baseline and Restaging (18)F-FDG PET Imaging Biomarkers in Patients with Esophageal Cancer. Radiology 2018; 287: 983–992. doi:10.1148/radiol.2018172229

[64] Rishi A, Zhang GG, Yuan Z et al. Pretreatment CT and (18) F-FDG PET-based radiomic model predicting pathological complete response and loco-regional control following neoadjuvant chemoradiation in oesophageal cancer. J Med Imaging Radiat Oncol 2021; 65: 102–111. doi:10.1111/1754-9485.13128

[65] Rahmim A, Bak-Fredslund KP, Ashrafinia S et al. Prognostic modeling for patients with colorectal liver metastases incorporating FDG PET radiomic features. Eur J Radiol 2019; 113: 101–109. doi:10.1016/j.ejrad.2019.02.006

[66] Yang L, Chu W, Li M et al. Radiomics in Gastric Cancer: First Clinical Investigation to Predict Lymph Vascular Invasion and Survival Outcome Using (18)F-FDG PET/CT Images. Front Oncol 2022; 12: 836098. doi:10.3389/fonc.2022.836098

[67] Atkinson C, Ganeshan B, Endozo R et al. Radiomics-Based Texture Analysis of (68)Ga-DOTATATE Positron Emission Tomography and Computed Tomography Images as a Prognostic Biomarker in Adults With Neuroendocrine Cancers Treated With (177)Lu-DOTATATE. Front Oncol 2021; 11: 686235. doi:10.3389/fonc.2021.686235

[68] Ma J, Guo D, Miao W et al. The value of (18)F-FDG PET/CT-based radiomics in predicting perineural invasion and outcome in non-metastatic colorectal cancer. Abdom Radiol (NY) 2022; 47: 1244–1254. doi:10.1007/s00261-022-03453-0

[69] Wei L, Cui C, Xu J et al. Tumor response prediction in (90)Y radioembolization with PET-based radiomics features and absorbed dose metrics. EJNMMI Phys 2020; 7: 74. doi:10.1186/s40658-020-00340-9

[70] Amyar A, Modzelewski R, Vera P et al. Weakly Supervised Tumor Detection in PET Using Class Response for Treatment Outcome Prediction. J Imaging 2022; 8. doi:10.3390/jimaging8050130

[71] Eertink JJ, van de Brug T, Wiegers SE et al. (18)F-FDG PET baseline radiomics features improve the prediction of treatment outcome in diffuse large B-cell lymphoma. Eur J Nucl Med Mol Imaging 2022; 49: 932–942. doi:10.1007/s00259-021-05480-3

[72] Parvez A, Tau N, Hussey D et al. (18)F-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 2018; 32: 410–416. doi:10.1007/s12149-018-1260-1

[73] Zhang X, Chen L, Jiang H et al. A novel analytic approach for outcome prediction in diffuse large B-cell lymphoma by [(18)F]FDG PET/CT. Eur J Nucl Med Mol Imaging 2022; 49: 1298–1310. doi:10.1007/s00259-021-05572-0

[74] Milgrom SA, Elhalawani H, Lee J et al. A PET Radiomics Model to Predict Refractory Mediastinal Hodgkin Lymphoma. Sci Rep 2019; 9: 1322. doi:10.1038/s41598-018-37197-z

[75] Ortega C, Eshet Y, Prica A et al. Combination of FDG PET/CT Radiomics and Clinical Parameters for Outcome Prediction in Patients with Hodgkin's Lymphoma. Cancers (Basel) 2023; 15. doi:10.3390/cancers15072056

[76] Eertink JJ, Zwezerijnen GJC, Cysouw MCF et al. Comparing lesion and feature selections to predict progression in newly diagnosed DLBCL patients with FDG PET/CT radiomics features. Eur J Nucl Med Mol Imaging 2022; 49: 4642–4651. doi:10.1007/s00259-022-05916-4

[77] Li M, Yao H, Zhang P et al. Development and validation of a [(18)F]FDG PET/CT-based radiomics nomogram to predict the prognostic risk of pretreatment diffuse large B cell lymphoma patients. Eur Radiol 2023; 33: 3354–3365. doi:10.1007/s00330-022-09301-5

[78] Frood R, Clark M, Burton C et al. Discovery of Pre-Treatment FDG PET/CT-Derived Radiomics-Based Models for Predicting Outcome in Diffuse Large B-Cell Lymphoma. Cancers (Basel) 2022; 14. doi:10.3390/cancers14071711

[79] Mazzara S, Travaini L, Botta F et al. Gene expression profiling and FDG-PET radiomics uncover radiometabolic signatures associated with outcome in DLBCL. Blood Adv 2023; 7: 630–643. doi:10.1182/bloodadvances.2022007825

[80] Ceriani L, Milan L, Cascione L et al. Generation and validation of a PET radiomics model that predicts survival in diffuse large B cell lymphoma treated with R-CHOP14: A SAKK 38/07 trial post-hoc analysis. Hematol Oncol 2022; 40: 11–21. doi:10.1002/hon.2935

[81] Chang CC, Chen CH, Hsieh JG et al. Iterated cross validation method for prediction of survival in diffuse large B-cell lymphoma for small size dataset. Sci Rep 2023; 13: 1438. doi:10.1038/s41598-023-28394-6

[82] 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 2019; 46: 2760–2769. doi:10.1007/s00259-019-04420-6

[83] Mu W, Liang Y, Hall LO et al. (18)F-FDG PET/CT Habitat Radiomics Predicts Outcome of Patients with Cervical Cancer Treated with Chemoradiotherapy. Radiol Artif Intell 2020; 2: e190218. doi:10.1148/ryai.2020190218

[84] Ferreira M, Lovinfosse P, Hermesse J et al. [(18)F]FDG PET radiomics to predict disease-free survival in cervical cancer: a multi-scanner/center study with external validation. Eur J Nucl Med Mol Imaging 2021; 48: 3432–3443. doi:10.1007/s00259-021-05303-5

[85] Nakajo M, Jinguji M, Tani A et al. Application of a Machine Learning Approach for the Analysis of Clinical and Radiomic Features of Pretreatment [(18)F]-FDG PET/CT to Predict Prognosis of Patients with Endometrial Cancer. Mol Imaging Biol 2021; 23: 756–765. doi:10.1007/s11307-021-01599-9

[86] de Alencar NRG, Machado MAD, Mourato FA et al. Exploratory analysis of radiomic as prognostic biomarkers in (18)F-FDG PET/CT scan in uterine cervical cancer. Front Med (Lausanne) 2022; 9: 1046551. doi:10.3389/fmed.2022.1046551

[87] Nakajo M, Jinguji M, Tani A et al. Machine learning based evaluation of clinical and pretreatment (18)F-FDG-PET/CT radiomic features to predict prognosis of cervical cancer patients. Abdom Radiol (NY) 2022; 47: 838–847. doi:10.1007/s00261-021-03350-y

[88] Lucia F, Visvikis D, Desseroit MC et al. Prediction of outcome using pretreatment (18)F-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2018; 45: 768–786. doi:10.1007/s00259-017-3898-7

[89] Bowen SR, Yuh WTC, Hippe DS et al. Tumor radiomic heterogeneity: Multiparametric functional imaging to characterize variability and predict response following cervical cancer radiation therapy. J Magn Reson Imaging 2018; 47: 1388–1396. doi:10.1002/jmri.25874

[90] Spohn SKB, Schmidt-Hegemann NS, Ruf J et al. Development of PSMA-PET-guided CT-based radiomic signature to predict biochemical recurrence after salvage radiotherapy. Eur J Nucl Med Mol Imaging 2023; 50: 2537–2547. doi:10.1007/s00259-023-06195-3

[91] Roll W, Schindler P, Masthoff M et al. Evaluation of (68)Ga-PSMA-11 PET-MRI in Patients with Advanced Prostate Cancer Receiving (177)Lu-PSMA-617 Therapy: A Radiomics Analysis. Cancers (Basel) 2021; 13. doi:10.3390/cancers13153849

[92] Assadi M, Manafi-Farid R, Jafari E et al. Predictive and prognostic potential of pretreatment (68)Ga-PSMA PET tumor heterogeneity index in patients with metastatic castration-resistant prostate cancer treated with 177Lu-PSMA. Front Oncol 2022; 12: 1066926. doi:10.3389/fonc.2022.1066926

[93] Alongi P, Stefano A, Comelli A et al. Radiomics analysis of 18F-Choline PET/CT in the prediction of disease outcome in high-risk prostate cancer: an explorative study on machine learning feature classification in 94 patients. Eur Radiol 2021; 31: 4595–4605. doi:10.1007/s00330-020-07617-8

[94] Tu SJ, Tran VT, Teo JM et al. Utility of radiomic zones for risk classification and clinical outcome predictions using supervised machine learning during simultaneous (11) C-choline PET/MRI acquisition in prostate cancer patients. Med Phys 2021; 48: 5192–5201. doi:10.1002/mp.15064

[95] Li P, Wang X, Xu C et al. 18)F-FDG PET/CT radiomic predictors of pathologic complete response (pCR. Eur J Nucl Med Mol Imaging 2020; 47: 1116–1126. doi:10.1007/s00259-020-04684-3

[96] Gómez OV, Herraiz JL, Udías JM et al. Analysis of Cross-Combinations of Feature Selection and Machine-Learning Classification Methods Based on [(18)F]F-FDG PET/CT Radiomic Features for Metabolic Response Prediction of Metastatic Breast Cancer Lesions. Cancers (Basel) 2022; 14. doi:10.3390/cancers14122922

[97] Choi JH, Kim HA, Kim W et al. Early prediction of neoadjuvant chemotherapy response for advanced breast cancer using PET/MRI image deep learning. Sci Rep 2020; 10: 21149. doi:10.1038/s41598-020-77875-5

[98] Stefano A, Comelli A, Bravatà V et al. A preliminary PET radiomics study of brain metastases using a fully automatic segmentation method. BMC Bioinformatics 2020; 21: 325. doi:10.1186/s12859-020-03647-7

[99] Muzi M, Wolsztynski E, Fink JR et al. Assessment of the Prognostic Value of Radiomic Features in (18)F-FMISO PET Imaging of Hypoxia in Post-surgery Brain Cancer Patients: Secondary Analysis of Imaging Data from a Single-Center Study and the Multicenter ACRIN 6684 Trial. Tomography 2020; 6: 14–22. doi:10.18383/j.tom.2019.00023

[100] Carles M, Popp I, Starke MM et al. FET-PET radiomics in recurrent glioblastoma: prognostic value for outcome after re-irradiation? Radiat Oncol 2021; 16: 46. doi:10.1186/s13014-020-01744-8

[101] Zhao W, Huang X, Wang G et al. PET/MR fusion texture analysis for the clinical outcome prediction in soft-tissue sarcoma. Cancer Imaging 2022; 22: 7. doi:10.1186/s40644-021-00438-y

[102] Kim J, Jeong SY, Kim BC et al. Prediction of Neoadjuvant Chemotherapy Response in Osteosarcoma Using Convolutional Neural Network of Tumor Center (18)F-FDG PET Images. Diagnostics (Basel) 2021; 11. doi:10.3390/diagnostics11111976

[103] Küstner T, Vogel J, Hepp T et al. Development of a Hybrid-Imaging-Based Prognostic Index for Metastasized-Melanoma Patients in Whole-Body 18F-FDG PET/CT and PET/MRI Data. Diagnostics (Basel) 2022; 12. doi:10.3390/diagnostics12092102

[104] Flaus A, Habouzit V, de Leiris N et al. Outcome Prediction at Patient Level Derived from Pre-Treatment 18F-FDG PET Due to Machine Learning in Metastatic Melanoma Treated with Anti-PD1 Treatment. Diagnostics (Basel) 2022; 12. doi:10.3390/diagnostics12020388

[105] Lee H, Hyun SH, Cho YS et al. Cluster analysis of autoencoder-extracted FDG PET/CT features identifies multiple myeloma patients with poor prognosis. Sci Rep 2023; 13: 7881. doi:10.1038/s41598-023-34653-3

[106] Jamet B, Morvan L, Nanni C et al. Random survival forest to predict transplant-eligible newly diagnosed multiple myeloma outcome including FDG-PET radiomics: a combined analysis of two independent prospective European trials. Eur J Nucl Med Mol Imaging 2021; 48: 1005–1015. doi:10.1007/s00259-020-05049-6

[107] Cheng NM, Hsieh CE, Fang YHD et al. Development and validation of a prognostic model incorporating [(18)F]FDG PET/CT radiomics for patients with minor salivary gland carcinoma. EJNMMI Res 2020; 10: 74. doi:10.1186/s13550-020-00631-3

[108] Pavic M, Bogowicz M, Kraft J et al. FDG PET versus CT radiomics to predict outcome in malignant pleural mesothelioma patients. EJNMMI Res 2020; 10: 81. doi:10.1186/s13550-020-00669-3