Received: 12 September 2022 | Accepted: 28 December 2022

DOI: 10.1002/pros.24483

ORIGINAL ARTICLE

The Prostate WILEY

Lack of repeatability of radiomic features derived from PET scans: Results from a ^{18}F-DCFPyL test–retest cohort

Rudolf A. Werner MD[1,2] | Bilêl Habacha[3] | Susanne Lütje MD[3] | Lena Bundschuh MSc[4] | Alekandser Kosmala MD[1] | Markus Essler MD[3] | Thorsten Derlin MD[5] | Takahiro Higuchi MD, PhD[1] | Constantin Lapa MD[4] | Andreas K. Buck MD[1] | Kenneth J. Pienta MD[6] | Martin A. Lodge PhD[2] | Mario A. Eisenberger MD[7] | Mark C. Markowski MD[7] | Martin G. Pomper MD[2,6,7] | Michael A. Gorin MD[8] | Eric C. Frey MD[9] | Steven P. Rowe MD, PhD[2,6,7] | Ralph A. Bundschuh MD, PhD[3,4]

[1]Department of Nuclear Medicine, University Hospital Würzburg, Würzburg, Germany

[2]The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[3]Department of Nuclear Medicine, University Hospital Bonn, Bonn, Germany

[4]Nuclear Medicine, Medical Faculty, University of Augsburg, Augsburg, Germany

[5]Department of Nuclear Medicine, Hannover Medical School, Hannover, Germany

[6]Department of Urology, The James Buchanan Brady Urological Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[7]Johns Hopkins University School of Medicine, Sidney Kimmel Comprehensive Cancer Center, Baltimore, Maryland, USA

[8]Milton and Carroll Petrie Department of Urology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[9]Radiopharmaceutical Imaging and Dosimetry, LLC, Baltimore, Maryland, USA

Correspondence
Ralph A. Bundschuh, MD, PhD, Nuclear Medicine, Medical Faculty, University of Augsburg, Stenglinstr. 2, 86156 Augsburg, Germany.
Email: ralph.alexander.bundschuh@med.uni-augsburg.de

Abstract

Objectives: PET-based radiomic metrics are increasingly utilized as predictive image biomarkers. However, the repeatability of radiomic features on PET has not been assessed in a test–retest setting. The prostate-specific membrane antigen-targeted compound ^{18}F-DCFPyL is a high-affinity, high-contrast PET agent that we utilized in a test-retest cohort of men with metastatic prostate cancer (PC).

Methods: Data of 21 patients enrolled in a prospective clinical trial with histologically proven PC underwent two ^{18}F-DCFPyL PET scans within 7 days, using identical acquisition and reconstruction parameters. Sites of disease were segmented and a set of 29 different radiomic parameters were assessed on both scans. We determined repeatability of quantification by using Pearson's correlations, within-subject coefficient of variation (wCOV), and Bland–Altman analysis.

Rudolf A. Werner, Bilêl Habacha, Steven P. Rowe, and Ralph A. Bundschuh contributed equally to this study.

Results: In total, 230 lesions (177 bone, 38 lymph nodes, 15 others) were assessed on both scans. For all investigated radiomic features, a broad range of inter-scan correlation was found (*r*, 0.07–0.95), with acceptable reproducibility for entropy and homogeneity (wCOV, 16.0% and 12.7%, respectively). On Bland–Altman analysis, no systematic increase or decrease between the scans was observed for either parameter (±1.96 SD: 1.07/−1.30, 0.23/−0.18, respectively). The remaining 27 tested radiomic metrics, however, achieved unacceptable high wCOV (≥21.7%).

Conclusion: Many common radiomic features derived from a test–retest PET study had poor repeatability. Only Entropy and homogeneity achieved good repeatability, supporting the notion that those image biomarkers may be incorporated in future clinical trials. Those radiomic features based on high frequency aspects of images appear to lack the repeatability on PET to justify further study.

KEYWORDS

^{18}F-DCFPyL, prostate cancer, PSMA, radiomics, repeatability, textural features

1 | INTRODUCTION

Image-derived, mathematically extracted features (radiomics) have gained interest in recent years, for example, for outcome prediction or lesion detectability on positron emission tomography (PET).[1,2] Given the increasing use of such parameters for clinical and scientific applications, we aimed to determine the repeatability of radiomics in men imaged with two near-term ^{18}F-DCFPyL PET scans in a prospective test–retest setting.

PET using prostate-specific membrane antigen (PSMA) is becoming increasingly important for diagnosis, treatment planning, and therapy monitoring in patients with metastatic prostate cancer (PC).[3–8] In this regard, repeatability of quantitative parameters is vital,[9–11] for example, in men scheduled for PSMA-targeted radioligand therapy (RLT) or other antitumor regimens.[12,13] Of note, previous studies using ^{68}Ga-labeled PSMA PET radiotracers have already reported on good repeatability for conventional PET parameters in sites of disease and normal tissue.[9] Relative to those first-generation PET radiotracers, however,^{18}F-labeled PSMA-targeted radiopharmaceuticals such as ^{18}F-DCFPyL (piflufolastat F18, PYLARIFY®) have demonstrated superior image quality and performance.[4,14] Regarding ^{18}F-DCFPYL, specifically, previous studies have already reported on robust repeatability in a test–retest setting utilizing conventional parameters, including maximum and mean standardized uptake values (SUV_{max}, SUV_{mean}) and PET-based tumor volume (PSMA-TV).[10,11]

2 | MATERIALS AND METHODS

Many of the clinical methods for this paper have been previously reported.[11]

This study was registered at ClinicalTrials.gov (NCT03793543), under the auspices of a United States Food and Drug Administration Investigational New Drug application (IND 121064). This prospective study was approved by the institutional review board at the Johns Hopkins Hospital (IRB00174393).[11]

2.1 | Patients

Twenty-one men with histological proven PC were enrolled (ClinicalTrial.gov identifier NCT03793543).[11] This single-center cohort has been reported previously, without investigating reproducibility of radiomic parameters.[11]

2.2 | Imaging

Two ^{18}F-DCFPyL PET/CT examinations were performed within 3.7 ± 3.0 days (range 1 to 7 days) to determine test–retest repeatability. Imaging protocols were not changed throughout the study. None of the patients received tumor-specific therapy between scans.[11]

PET/CT examinations were performed on a Siemens Biograph 128 mCT (Siemens Healthineers). PET data was acquired (3 min per bed position) approximately 60 min after the injection of ^{18}F-DCFPyL. A low dose CT ((no contrast, 120 kV, 40 effective mAs, 0.5 tube rotation time, and pitch of 0.8) was acquired for attenuation and scatter correction and anatomic localization. PET data was reconstructed using a standard ordered-subset expectation maximization algorithm (three-dimensional ordered-subsets expectation-maximization with time-of-flight, 2 iterations 21 subsets, Gaussian filter with 5 mm full-width-at-half-maximum,

TABLE 1 Radiomic parameters for all lesions ($n = 230$) with Pearson correlation, Kendall's τ and within-subject coefficient of variation (wCOV)

	r	Kendall's τ	wCOV [%]
Busyness	0.70	0.67	21.7
Coarsness	0.67	0.67	31.6
Contrast	0.92	0.76	90.8
Entropy	0.95	0.78	16.0
GLNU	0.79	0.70	33.9
HGLRE	0.93	0.81	68.2
HGLZE	0.93	0.78	69.9
Homogeneity	0.90	0.75	12.7
Intensity variation	0.39	0.30	31.3
Kurtosis	0.70	0.36	84.0
LRE	0.63	0.63	55.7
LRHGLE	0.92	0.60	40.6
LRLGLE	0.66	0.75	119.7
LZE	0.28	0.63	630.4
LZHGLE	0.14	0.49	511.3
LZLGLE	0.32	0.72	873.6
LGLRE	0.75	0.76	63.7
LGLZE	0.57	0.73	82.6
RLNU	0.58	0.52	28.6
Run percentage	0.44	0.36	28.6
SRE	0.70	0.50	27.4
SRHGLE	0.92	0.73	82.4
SRLGLE	0.42	0.52	101.3
SZE	0.62	0.49	30.6
SZHGLE	0.92	0.73	76.2
SZLGLE	0.07	0.32	161.6
Size variation	0.79	0.70	33.9
ZLNU	0.39	0.30	31.3
Zone percentage	0.53	0.43	45.5

Note: Only entropy and homogeneity (marked in bold) achieved wCOVs below 20%, thereby indicating good reproducibility.

Abbreviations: GLNU, Gray Level Non Uniformity; HGLRE, High Gray Level Run Emphasis; HGLZE, High Gray Level Zone Emphasis; LGLRE, Low Gray Level Run Emphasis; LGLZE, Low Gray Level Zone Emphasis; LRE, Long Run Emphasis; LRHGLE, Long Run High Gray Level Emphasis; LRLGLE, Long Run Low Gray Level Emphasis; LZE, Long Zone Emphasis; LZHGLE, Long Zone High Gray Level Emphasis; LZLGLE, Long Zone Low Gray Level Emphasis; RLNU, Run Length Non Uniformity; SRE, Short Run Emphasis; SRHGLE, Short Run High Gray Level Emphasis; SRLGLE, Short Run Low Gray Level Emphasis; SZE, Short Zone Emphasis; SZHGLE, Short Zone High Gray Level Emphasis; SZLGLE, Short Zone Low Gray Level Emphasis; ZLNU, Zone Length Non Uniformity.

4 mm voxel size) implemented by the manufacturer of the scanner including attenuation and scatter correction.[18]F-DCFPyL was synthesized according to current good manufacturing practices as previously described.[15] More details on image protocols can be found in.[11]

2.3 | Image analysis

As described in,[11] the segmentation of tumor locations was performed in a consensus review by three nuclear medicine physicians with experience in reading PSMA-targeted PET (BH,

RAB, RAW). We used InterView Fusion software (Version 3.08.005.0000, Mediso Medical Imaging Ltd) for image analysis. The entire volume of all ^{18}F-DCFPyL-avid tumor lesions was manually segmented.[11] Within the segmented volume normalization and discretization of the activity values was performed to a total of 64 bins. Consequently, within the segmented volume of interest, 29 radiomic parameters were calculated as automatically provided by the applied software (Supporting Information: Table 1). Derived metrics are listed in Table 1 and further details can be found in.[16,17]

2.4 | Statistical analysis

Evaluated radiomics features were individually compared between the two scans. Details on statistics can also be found in.[11] We plotted scatter diagrams and linear regression analysis for all lesions. Bland–Altman plots were calculated for the absolute and relative differences of these data and upper and lower level of agreement were also recorded.[18,19] Pearson's correlation was performed (providing *r*). In addition, Kendall´s tau (τ) was calculated.[20,21] The within-subject coefficient of variation (wCOV) was also assessed according to Synek.[22] The wCOV was estimated for subgroups of the lesions (LN, osseous, and other lesions). For comparison of different wCOVs, the method of Forkmann[23] was used. We also calculated the dice coefficient which allowed us to assess the similarity of respective volumes of interests (VOIs). The dice coefficient is defined as two times the volume of the intersection of the two VOIs divided by the sum of the two volumes. Ranging from 0, no similarity to 1, similar, a dice score of minimum 0.7 was considered acceptable as done before.[24] Calculation of the dice score was performed using the VOIs obtained in InterviewFusion software (Mediso Medical Systems Ltd., Budapest, Hungary). A $p < 0.05$ was considered significant. We used MedCalc software package (Version 19.6, MedCalc software Ltd) and Microsoft Excel 2016 (Microsoft Cooperation).

3 | RESULTS

Prostate-specific antigen levels at time of scan were 22.3 ± 34.3, ranging from 0.4 to 138.4 ng/ml.[11] For further details including inclusion and exclusion criteria and patient's characteristics (age, height, weight, and therapies before imaging), please refer to.[11]

A total of 230 lesions (177 bone, 38 lymph nodes (LN), 15 others) were delineated on both scans. Serving as quality control, although limited due to potential errors in co-registration of the two image data sets, dice coefficiant on test and retest VOIs was in mean 0.48 (range 0.03–0.99), thereby considered acceptable. We observed a broad range of inter-scan correlation among all investigated parameters, ranging from 0.07 (short zone low gray level emphasis) to 0.95 (entropy). With Kendalls's τ, comparable findings were recorded, ranging from 0.30 for intensity variation up to 0.81 for high gray level run emphasis. Analyzing all lesions, the lowest wCOVs indicating excellent reproducibility were found for entropy and homogeneity (wCOV, 16.0% and 12.7%, respectively). The remaining 27 out of 29 (93.1%) investigated radiomic features, however, demonstrated unaccaptably high wCOV, with up to 873.6% for long tone low gray level emphasis. *R*-values, τ, and wCOV can be found in Table 1. For entropy and homogeneity, Figure 1 provides scatter plots (top), and Bland–Altman Plots for absolute (middle) and relative differences (bottom). No systematic increase or decrease between the scans was observed for either parameter (±1.96 SD: 1.07/−1.30, 0.23/−0.18, respectively; Figure 1, middle). On a lesion-type level, radiomic features derived from LN and bone metastases, however, did not show any significant differences (wCOV entropy: LN, 16.4% vs. skeleton, 14.9%, $p = 0.42$; wCOV homogeneity: LN, 14.2% vs. skeleton, 12.5%, $p = 0.26$). Visceral lesions were not analyzed further, as the overall number of lesions attributable to this organ compartment was too low. Other textural parameters showed statistical significant different wCOV for lymph nodes and bone lesion, however, all of them showed much higher wCOV that entropy and homogeneity. Supporting Information: Table 2 provides an overview of wCOV on an organ-based level, including LN and osseous lesions, for all other investigated radiomic features.

4 | DISCUSSION

Our work has shown that the majority of radiomic features derived from PET images are not repeatable. A potential explanation of the lack of repeatability of high-frequency radiomics parameters is the fact that voxel values in PET are not estimable parameters.[25] This means that unbiased estimators for voxels do not exist, and that voxel values can be corrupted by arbitrary amounts of null functions of the imaging system. In PET, null functions are activity distributions that have discrete projections equal to zero in the absence of noise, and correspond to the high-frequency, fine-detail features of the activity distributions. Since the projections do not provide information about magnitude of these null function components, arbitrary amounts of them can be introduced by the reconstruction process. This is exacerbated by the noise in the projection data and the fact that representing the continuous activity distribution in the patient by a voxelized representation results in modeling errors. The net result is that high-frequency features of the image are likely unreliable. This is a reason that such features are conventionally controlled using low-pass filters or Bayesian reconstruction methods. The net result is that radiomics features that make use of high-frequency features of voxelized PET images will likely contain null function components. Since images from different measurements can contain arbitrary amounts of these null function components, this has the potential to lead to variability of the values from repeated measurements.

Based on a ^{18}F-DCFPyL PET/CT prospective cohort in the test-retesting setting, we observed high repeatability of entropy and homogeneity, providing wCOV values below 20% and correlative indices of ≥0.90. In a previous study analyzing conventional PET parameters in the identical cohort, such high correlation had only been

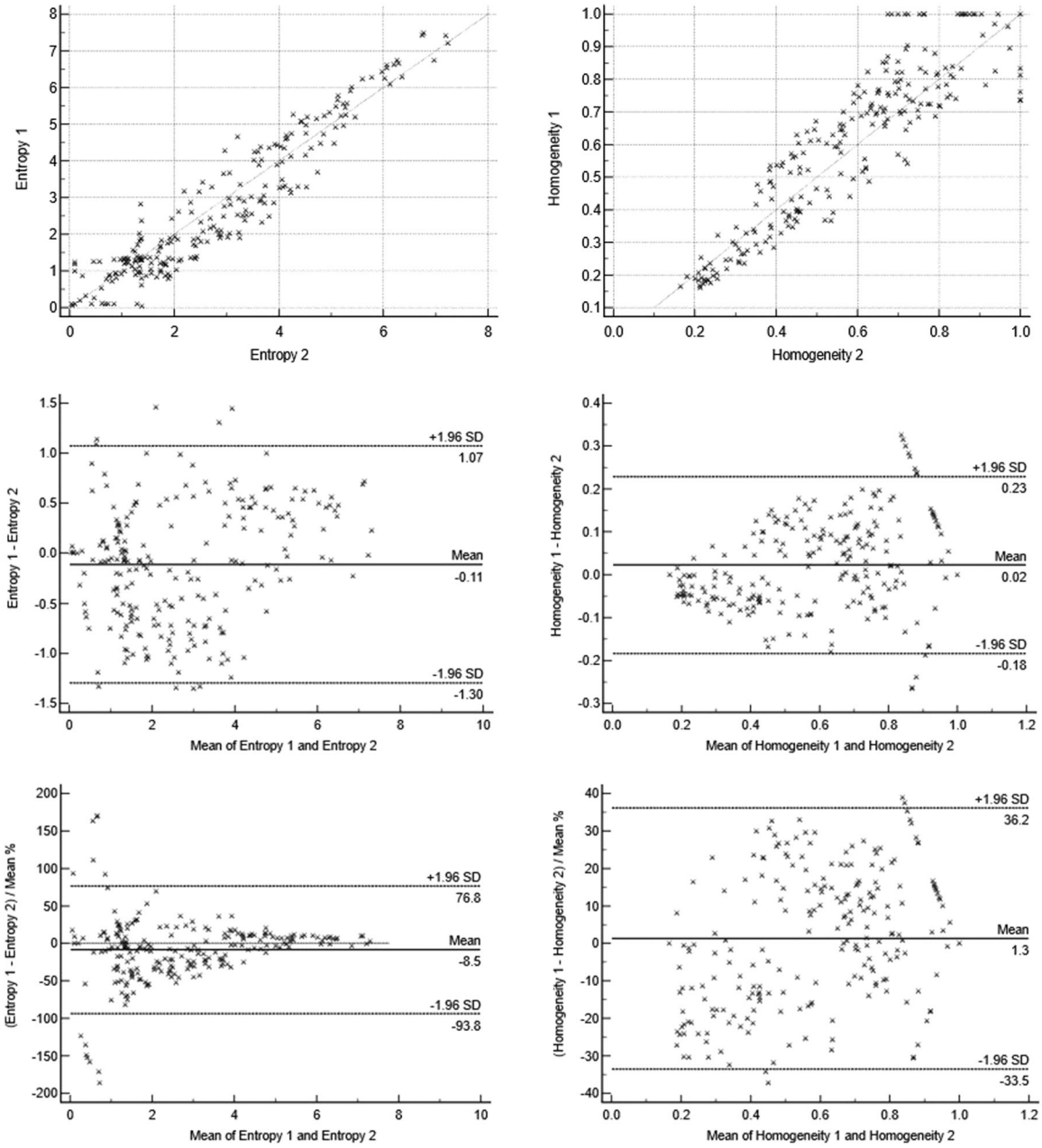

FIGURE 1 Correlations (upper rows), Bland–Altman-plots for absolute (middle rows) and relative differences (bottom rows) for the radiomic parameters Entropy (left) and Homogeneity (right). While good correlations were observed for both parameters, higher Entropy values had more robust repeatability, in particular for relative values (bottom left).

recorded for SUV_{max} and SUV_{mean}, supporting the notion that SUVs, entropy, and homogeneity may serve as reliable and reproducible quantitative parameters on ^{18}F-DCFPyL PET.[11] However, those radiomic features are lesion-level and not subject to the high-frequency variability of most of the described radiomic features, including "high-level features" such as gray tone difference-derived radiomics features. Those findings may be of importance for radiomics analyses that will be tied to relevant clinical endpoints, for example, to determine the performance of PET-based radiomics to predict biochemical response or overall survival in men treated with PSMA-directed RLT.[26,27] For high-frequency radiomic features, which included 27/29 (93%) of the herein investigated radiomic parameters, reproducibility was low, with unacceptably high wCOV of up to 873.6%.

^{18}F-DCFPyL was the first commercial PSMA-ligand that was FDA approved for imaging men with PC[28] and thus, an even more widespread adoption of this PET agent can be assumed not only in the clinic, but also in major trials. Most likely, quantitative assessments of obtained PET data will not be limited to conventional parameters (SUV, PSMA-TV), but also include more sophisticated image analysis, for example, mathematically extracted textural parameters such as radiomics.[26,27] In our previous study exclusively focusing on standard parameters using ^{18}F-DCFPyL,[11] SUVs

achieved acceptable reproducibility, ranging from 7.3% for SUV_{mean} to 12.1% for SUV_{max}. In the identical cohort, among 29 radiomic parameters tested in the present investigation, comparable low wCOVs were only recorded for the radiomic features entropy and homogeneity (16.0% and 12.7%, respectively).

Of note, regardless of which statistical test (Kendall's τ, Pearson's correlation) was applied, correlative indices for the latter parameters remained high, even in a subanalysis focusing on either LN or skeletal lesions. As such, given previous[11] and present findings, clinical trials using ^{18}F-DCFPyL should focus on SUVs, with lesion-level radiomic parameters such as entropy and homogeneity as repeatable imaging parameters. All other radiomic parameters, should not be used for quantification, as for those metrics, wCOV ranged from 21.7% to 873.6% (Table 1), suggesting unacceptable reproducibility. Those findings are in line with a recent meta-analysis for PET imaging (mainly on ^{18}F-FDG), demonstrating that zone matrix- and gray tone difference-derived radiomics should be used with caution, as they are subject to an inherent high sensivitiy to variation. This is in contrast to conventional markers such as SUV, which have been considered robust in the context of PET-based quantification.[29] In this regard, acquisition parameters, including spatial resolution, reconstruction type, and discretization of gray-level intensities have been advocated to have a relevant impact on feature robustness.[29,30] Although a bias caused by those factors cannot be completely ruled out in our study, the identical PET scanner with the same reconstruction algorithm and the same software tool for image analyses have been used. Thus, technical considerations have not caused the observed unacceptably low reproducibility of many of the parameters. In addition, a biology-driven impact can also be ruled out, as none of the included patients received antitumor treatment between scans and substantial disease progression is unlikely to occur within such a short time interval of repeated ^{18}F-DCFPyL PETs (maximum, 7 days). As such, similar to reproducibility assessments in the context of ^{18}F-FDG for oncolocy imaging,[29] we provide evidence that the majority of the radiomic biomarkers investigated are not applicable for PET.

Of note, for relative values of entropy, we observed that higher values had more robust repeatability (Figure 1, bottom left). Similar findings have also been recorded for SUV_{mean} and SUV_{max}.[11] Such observations may be of relevance in the context of outcome studies or when assessing the percentage change of those parameters between baseline and follow-up PSMA PET.[12] In this regard, the delta of ^{18}F-DCFPyL PET-derived SUV_{max} was linked to clinical outcome in men after initiating abiraterone or enzelutamide, including survival and time to novel treatment.[12] Although such studies are not available to date for radiomics and antiandrogen treatment, future studies may focus on entropy to investigate the predictive potential or associations between delta-entropy and clinical endpoints, preferably in men under such common therapeutic regimens. In metastatic castration-resistant PC patients, Moazemi et al.[1] have already reported on radiomics signatures and various clinical parameters for predicting survival in patients scheduled for PSMA-targeted RLT. Relative to the present study, however, a first-generation ^{68}Ga-labeled PET agent was applied[1] and thus, for ^{18}F-labeled PET, we have now provided a roadmap as to which radiomic parameters should be included for such an outcome analysis. In addition, Zamboglou and colleagues recently reported on the usefulness of radiomics to discriminate between patients with intermediate versus high-risk PC by comparing their PET-based results with ex-vivo specimen. Again, ^{68}Ga-labeled PSMA PET was used.[31] Thus, while the study design provided by Zamboglou et al.[31] may serve as a template, the herein provided findings may assist in selecting the appropriate radiomic metrics if ^{18}F-DCFPyL will be applied to similar studies. Further, with the risk of false discovery that is inherent to radiomics studies with numerous parameters, we suggest that radiomic studies with PET be undertaken with significant caution and that the field should focus on lesion-level characteristics and abandon features based on fine detail in the images. Nonetheless, future investigations should also determine whether the herein provided wCOV would be also recorded when other software tools or different radiotracers are used, for example,^{68}Ga-labeled PSMA PETs.[32]

5 | CONCLUSION

In this PET study, among the 29 investigated radiomic features on both scans, we found good repeatability only for the parameters entropy and homogeneity, leaving 27/29 (93%) of the putative imaging biomarkers with an unacceptably low repeatability. Of note, no significant differences in wCOV were observed for lymphatic or osseous disease spread for entropy and homogeneity, suggesting that such radiomic features are highly reproducible even among different organ compartments. As such, those radiomic image biomarkers may be subject of future studies or clinical trials, while other parameters may not add additional information for PSMA-directed PET using ^{18}F-DCFPyL. However, unnacceptably high wCOVs were observed with all other radiomic features, suggesting that radiomics of PET must be limited low-frequency features and exclude high-frequency, fine-detail features. The entire field of PET radiomics should be geared to more meaningful parameters.

AUTHOR CONTRIBUTIONS

Rudolf A. Werner, Bilêl Habacha, Thorsten Derlin, Martin G. Pomper, Michael A. Gorin, Steven P. Rowe, and Andreas K. Buck designed the study, wrote the manuscript, and researched data. Rudolf A. Werner, Bilêl Habacha, Lena Bundschuh, Steven P. Rowe, and Andreas K. Buck performed analysis. All authors aided in drafting the manuscript and revised it critically for important intellectual content. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

Funding for this study was received from the Movember Foundation, the Prostate Cancer Foundation Young Investigator Award and National Institutes of Health grants CA134675, CA184228, EB024495, and CA183031. Open Access funding enabled and organized by Projekt DEAL.

CONFLICTS OF INTEREST

M. G. P. is a co-inventor on a US patent covering ^{18}F-DCFPyL and as such is entitled to a portion of any licensing fees and royalties generated by this technology. This arrangement has been reviewed and approved by the Johns Hopkins University in accordance with its conflict of interest policies. S. P. R. is a consultant for Progenics Pharmaceuticals Inc., the licensee of ^{18}F-DCFPyL. R. A. B. is Consultant for Bayer Healthcare (Leverkusen, Germany) and Eisai GmbH (Frankfurt, Germany). R. A. B. has a noncommercial research agreement and is on the speakers list with Mediso Medical Imaging (Budapest, Hungary). M. E. is Consultant for Bayer Healthcare (Leverkusen, Germany) and Eisai GmbH (Frankfurt, Germany), IPSEN and Novartis. The result of this work is part of the doctoral thesis of B. H., planned to be submitted at the Medical Faculty of the University of Bonn. The remaining authors declare no conflict of interest as well as consent for scientific analysis and publication.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, to any qualified researcher.

ORCID

Thorsten Derlin http://orcid.org/0000-0002-8786-4105
Mark C. Markowski http://orcid.org/0000-0003-2780-5100
Steven P. Rowe http://orcid.org/0000-0003-2897-4694
Ralph A. Bundschuh https://orcid.org/0000-0001-7291-9372

REFERENCES

1. Moazemi S, Erle A, Lütje S, Gaertner FC, Essler M, Bundschuh RA. Estimating the potential of radiomics features and radiomics signature from pretherapeutic PSMA-PET-CT scans and clinical data for prediction of overall survival when treated with (177)Lu-PSMA. *Diagnostics*. 2021;11(2):186.
2. Zamboglou C, Bettermann AS, Gratzke C, et al. Uncovering the invisible-prevalence, characteristics, and radiomics feature-based detection of visually undetectable intraprostatic tumor lesions in (68)GaPSMA-11 PET images of patients with primary prostate cancer. *Eur J Nucl Med Mol Imaging*. 2021;48(6):1987-1997.
3. Choyke PL, Bouchelouche K. Prostate specific membrane antigen (PSMA) imaging: the past is prologue. *Transl Androl Urol*. 2019;8(4):283-285.
4. Werner RA, Derlin T, Lapa C, et al. ^{18}F-labeled, PSMA-Targeted radiotracers: leveraging the advantages of radiofluorination for prostate cancer molecular imaging. *Theranostics*. 2020;10(1):1-16.
5. Miyahira AK, Pienta KJ, Babich JW, et al. Meeting report from the Prostate Cancer Foundation PSMA theranostics state of the science meeting. *Prostate*. 2020;80(15):1273-1296.
6. Habl G, Sauter K, Schiller K, et al. (68) Ga-PSMA-PET for radiation treatment planning in prostate cancer recurrences after surgery: individualized medicine or new standard in salvage treatment. *Prostate*. 2017;77(8):920-927.
7. Hope TA, Eiber M, Armstrong WR, et al. Diagnostic accuracy of 68Ga-PSMA-11 PET for pelvic nodal metastasis detection prior to radical prostatectomy and pelvic lymph node dissection: a multicenter prospective phase 3 imaging trial. *JAMA Oncol*. 2021;7(11):1635-1642.
8. Seifert R, Seitzer K, Herrmann K, et al. Analysis of PSMA expression and outcome in patients with advanced Prostate Cancer receiving (177)Lu-PSMA-617 Radioligand Therapy. *Theranostics*. 2020;10(17):7812-7820.
9. Pollard JH, Raman C, Zakharia Y, et al. Quantitative test-retest measurement of (68)Ga-PSMA-HBED-CC in tumor and normal tissue. *J Nucl Med*. 2020;61(8):1145-1152.
10. Jansen BHE, Cysouw MCF, Vis AN, et al. Repeatability of quantitative (18)F-DCFPyL PET/CT measurements in metastatic prostate cancer. *J Nucl Med*. 2020;61(9):1320-1325.
11. Werner RA, Habacha B, Lütje S, et al. High SUVs have more robust repeatability in patients with metastatic prostate cancer: results from a prospective test-retest cohort imaged with (18)F-DCFPyL. *Mol Imaging*. 2022;2022:1-10.
12. Zukotynski KA, Emmenegger U, Hotte S, et al. Prospective, single-arm trial evaluating changes in uptake patterns on prostate-specific membrane antigen-targeted (18)F-DCFPyL PET/CT in patients with castration-resistant prostate cancer starting abiraterone or enzalutamide. *J Nucl Med*. 2021;62(10):1430-1437.
13. Hartrampf PE, Weinzierl FX, Buck AK, et al. Matched-pair analysis of [(177)Lu]Lu-PSMA I&T and [(177)Lu]Lu-PSMA-617 in patients with metastatic castration-resistant prostate cancer. *Eur J Nucl Med Mol Imaging*. 2022;49(9):3269-3276.
14. Dietlein M, Kobe C, Kuhnert G, et al. Comparison of [(18)F]DCFPyL and [(68)Ga]Ga-PSMA-HBED-CC for PSMA-PET imaging in patients with relapsed prostate cancer. *Mol Imaging Biol*. 2015;17(4):575-584.
15. Szabo Z, Mena E, Rowe SP, et al. Initial evaluation of [(18)F]DCFPyL for prostate-specific membrane antigen (PSMA)-targeted PET imaging of prostate cancer. *Mol Imaging Biol*. 2015;17(4):565-574.
16. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-762.
17. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563-577.
18. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327(8476):307-310.
19. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound Obstet Gynecol*. 2003;22(1):85-93.
20. Rupinski MT, Dunlap WP. Approximating Pearson product-moment correlations from Kendall's tau and Spearman's rho. *Educ Psychol Meas*. 1996;56(3):419-429.
21. Walker DA. MASM9: converting Kendall's tau for correlational or meta-analytic analyses. *J Mod Appl Stat Methods*. 2003;2(2):525-530.
22. Synek V. Evaluation of the standard deviation from duplicate results. *Accreditation Qual Assur*. 2008;13:335-337.
23. Forkmann J. Stimator and tests for common coefficients of variation in normal distributions. *Commun. Stat Theory Methods*. 2009;38:233-251.
24. Baskaran L, Al'Aref SJ, Maliakal G, et al. Automatic segmentation of multiple cardiovascular structures from cardiac computed tomography angiography images using deep learning. *PLoS One*. 2020;15(5):e0232573.
25. Barrett HH, Myers KJ, eds. Signal recovery as estimation: a discourse on null functions and nuisance parameters. *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings on CD-ROM*. Optica Publishing Group; 2007.
26. 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*. 2021;13(15):3849.
27. Khurshid Z, Ahmadzadehfar H, Gaertner FC, et al. Role of textural heterogeneity parameters in patient selection for 177Lu-PSMA therapy via response prediction. *Oncotarget*. 2018;9(70):33312-33321.
28. FDA. Approves (18)F-DCFPyL PET agent in prostate cancer. *J Nucl Med*. 2021;62(8):11N.

29. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging*. 2019;46(13):2638-2655.
30. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to radiomics. *J Nucl Med*. 2020;61(4):488-495.
31. Zamboglou C, Carles M, Fechter T, et al. Radiomic features from PSMA PET for non-invasive intraprostatic tumor discrimination and characterization in patients with intermediate- and high-risk prostate cancer—a comparison study with histology reference. *Theranostics*. 2019;9(9):2595-2605.
32. Zang S, Ai S, Yang R, et al. Development and validation of (68)Ga-PSMA-11 PET/CT-based radiomics model to detect primary prostate cancer. *EJNMMI Res*. 2022;12(1):63.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Werner RA, Habacha B, Lütje S, et al. Lack of repeatability of radiomic features derived from PET scans: results from a ^{18}F-DCFPyL test–retest cohort. *Prostate*. 2023;1-8. doi:10.1002/pros.24483