Computers in Biology and Medicine 133 (2021) 104400

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

The impact of the variation of imaging parameters on the robustness of Computed Tomography radiomic features: A review

Reza Reiazi [a,b,c], Engy Abbas [d], Petra Famiyeh [a,e], Aria Rezaie [a], Jennifer Y.Y. Kwan [a], Tirth Patel [i], Scott V. Bratman [a,c], Tony Tadic [a,c], Fei-Fei Liu [a,c], Benjamin Haibe-Kains [a,b,f,g,h,*]

[a] Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada
[b] Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada
[c] Department of Radiation Oncology, University of Toronto, Toronto, Ontario, Canada
[d] Joint Department of Medical Imaging, University of Toronto, Toronto, Canada
[e] Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada
[f] Ontario Institute for Cancer Research, Toronto, Ontario, Canada
[g] Department of Computer Science, University of Toronto, Toronto, Ontario, Canada
[h] Vector Institute, Toronto, Ontario, Canada
[i] Techna Institute, University Health Network, Toronto, Ontario, Canada

ARTICLE INFO

Keywords:
Radiomics
Robustness
Reproducibility
Imaging Parameters
CT Scan

ABSTRACT

The field of radiomics is at the forefront of personalized medicine. However, there is concern that high variation in imaging parameters will impact robustness of radiomic features and subsequently the performance of the predictive models built upon them. Therefore, our review aims to evaluate the impact of imaging parameters on the robustness of radiomic features. We also provide insights into the validity and discrepancy of different methodologies applied to investigate the robustness of radiomic features. We selected 47 papers based on our predefined inclusion criteria and grouped these papers by the imaging parameter under investigation: (i) scanner parameters, (ii) acquisition parameters and (iii) reconstruction parameters. Our review highlighted that most of the imaging parameters are disruptive parameters, and shape along with First order statistics were reported as the most robust radiomic features against variation in imaging parameters. This review identified inconsistencies related to the methodology of the reviewed studies such as the metrics used for robustness, the feature extraction techniques, the reporting style, and their outcome inclusion. We hope this review will aid the scientific community in conducting research in a way that is more reproducible and avoids the pitfalls of previous analyses.

1. Introduction

Computed Tomography (CT) is the modality of choice for the depiction, diagnosis and monitoring of many diseases. The consistency with which CT provides high-resolution images has led to its extended use in diagnostic, prognostic, quality assessment, and dose calculation in radiotherapy [1]. Moreover, the number of possible applications continues to grow because of the innovative approaches researchers have developed to extract new, potentially clinically-relevant features from radiological images [2].

Advances in the field of artificial intelligence introduced decision support based on quantitative image descriptors called radiomic features, as a new tool to assess medical images beyond the narrow visual inspectors. The main idea behind this new research field, called radiomics, is that advanced analysis of images can noninvasively amplify clinical prognostic nomographs, correlate imaging phenotypes with genomic and proteomic signatures, and subsequently improve clinical decision making. Although radiomics is currently limited in its applicability to daily clinical practice, radiomics signatures have increasingly become a critical component of precision medicine. The integration of image-derived data with patient-tailored diagnostic or prognostic methods may result in more personalized treatments in the near future [3,4].

As the radiomics field matured, it became apparent that the main drawback of radiomic features is their low robustness to variation in acquisition and reconstruction parameters, which may negatively affect the generalizability of models built upon those features [5–10]. The effect of variation of image acquisition on the robustness of radiomic

* Corresponding author. Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada.
E-mail address: benjamin.haibe-kains@uhnresearch.ca (B. Haibe-Kains).

https://doi.org/10.1016/j.compbiomed.2021.104400
Received 15 December 2020; Received in revised form 23 March 2021; Accepted 11 April 2021
Available online 16 April 2021

features has been found to be greater than that of segmentation [10] and interobserver variability [11]. This variation affects the information being extracted by image feature algorithms, which in turn affects classifier performance and is of paramount importance to ensure the successful application of CT-derived radiomics in the field of oncology [12,13]. Consequently, we must treat the reported performance of radiomics-based models with caution [14] as quantitative changes may be primarily due to acquisition variability rather than from real physio-pathological effects [15]. CT-derived radiomic features have intrinsic dependencies on voxel size and number of gray levels. Such dependencies demonstrate that their application is highly dependent on the careful selection of a nominal voxel size and the number of intensity bins [16] [(1)], [17], [18]. For instance, the type of binning and number of bins significantly affected radiomic parameters extracted from coronary artery plaques [19]. In other words, post-processing settings, such as bin size in intensity normalization or voxel size in voxel rescaling, should be adjusted based on the type of radiomic feature, imaging parameter, organ and clinical outcome. In addition, the numerical values of radiomic features were highly correlated with tumor volume and voxel resampling was not sufficient to remove this correlation [16]. One can eliminate this dependency only by including the number of voxels in feature definitions (i.e., feature normalization) [12]. One possible solution would be to focus on the imaging parameters that disturb the robustness of radiomic features. We refer to these imaging parameters as disruptive parameters (DP). Having known these parameters, we must select the radiomics features robust to them. However, this action requires deep knowledge about these DPs and their related robust features. In this review, we collected a list of disruptive (DP) and non-disruptive (nDP) parameters associated with radiomic features extracted from CT images.

2. Materials and methods

2.1. Literature search

Our literature search was conducted using PubMed and included all publications from February 2021 or earlier. Publications included in our review met all of the following inclusion criteria: (1) publications must be peer-reviewed English full-text reports; (2) have radiomics features extracted from CT images; and (3) focus on the robustness of radiomic features resulting from variation in scanner manufacturer, imaging acquisition or reconstruction parameters. We used the following search words: (Radiomics OR Texture) AND (Reproducibility OR Robustness OR Stability) AND (CT OR CT Scan OR Computed Tomography) specified in the search string. Furthermore, we excluded repeatability or test-retest studies. We also screened the Cochrane Database of Systematic Reviews for any previous systematic reviews addressing the robustness of CT-based radiomic features. For all the articles obtained where we used the full text for data extraction, we screened the bibliographic references within them for potentially eligible studies. The researchers downloaded these electronic full-text articles using university library subscriptions.

2.2. Data extraction

We extracted information including the subject (patient or phantom), the type of organ, the clinical implications for human studies and the type of phantom for phantom studies. We noted the study type (retrospective or prospective) as well as any image acquisition and reconstruction parameters explicitly stated in the text. We also recorded the total number of radiomic features tested and reported these features according to the following well known feature groups: Shape features, First Order Statistics (FO), higher-order textural features (GLCM, RLM, GLSZM, GLDZM, NGTDM, NGLDM). A complete list of the abbreviations used in this article has been presented in Supplementary Table 1. We also classified features based on 2D or 3D extraction. In addition, we also recorded the details of the software used to quantitatively extract radiomic features, as well as the statistical method(s) used to report feature robustness. Reporting complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards [20].

3. Results

The PubMed search yielded 206 abstracts, including 31 eligible studies that reported the impact of imaging parameters on the robustness of radiomic features extracted from CT images. We also included the full text for 16 papers located in the references of retrieved studies (Fig. 1). Finally, we derived in-depth evaluation on 47 studies, of which 30 human studies and 15 phantom studies and 2 studies worked on both patient and phantom data. The vast majority of human studies addressed the robustness of CT-based radiomic features in lungs (51.3%) (Fig. 2A). All prospective studies used different types of physical phantoms. Amongst the prospective phantom studies, Credence Cartridge Radiomics (CCR) Phantom developed by Mackin [14] was the most common type of phantom used (41.2%) (Fig. 2B).

3.1. Disruptive parameters

We grouped the disruptive parameters into 3 classes: (i) scanner (studies that only focused on the effect of different CT scanners disregarding any specific acquisition or reconstruction parameter); (ii) acquisition (including patient-related parameters); and (iii) reconstruction.

Scanner was evaluated in 10 studies (7 phantom and 3 patient studies) (Supplementary Table 2). In these studies different scanners with various image acquisition and reconstruction settings were used to mimic routine imaging protocols [14]. In 8 [10,14,17,21–25] studies, applying different feature extraction software, evaluation metrics and features, CT scanner was found as a disruptive parameter (Fig. 3). It is noteworthy to mention that different CT scanners have been proven to have variation in their Hounsfield units even with the same acquisition parameters [22,26]. Perrin et al. showed that utilizing images from different scanners reduced the number of liver tumor-derived robust features (CCC>0.9) from 75 to 35 (out of 254) [27]. However, Mackin et al. reported that variations of radiomic features among different scanners were found to be similar to their variation among 20 Non-small-cell lung carcinoma (NSCLC) patients stating Business-NGTDM and strength-NGTDM the most and least robust features respectively [14].

Acquisition parameters including exposure, respiratory motion, contrast agent administration, tube voltage and pitch were evaluated in 34 studies using phantom (13), patient (19) or both types of data (2) either alone or in combination with other parameters (i.e. reconstruction and scanner) (Supplementary Table 2). Exposure in 9 [9,21,23,24,28–32] vs. 6 [10,17,26,33–35], respiratory motion in 6 [36–41] vs none, contrast agent in 4 [10,27,34,42] vs. 1 [43] and tube voltage in 3 [21,23,24]vs. none were respectively reported as the most and least disruptive parameters for radiomic features (Fig. 3). Pitch was the only image acquisition parameter that was reported as a nDPs in all the corresponding studies [24,26,36]. In some studies, respiratory motion was not explicitly evaluated and instead the superiority of using average intensity projection images for radiomic studies over those taken with free-breathing has been shown [40]. Using scans at the end of the exhalation phase of a 4DCT acquisition [11] has been shown to yield more robust radiomics features (CCC $>$ 0.85). Another study stated the considerable impact of respiration on feature robustness and they suggested applying 4D stability as a feature selection to reduce feature variability [37]. The reviewed studies reported Shape features [37,40] and Short-Run Emphasis [39] as the Radiomic features most robust to respiration. Although 4D features were more robust, there was no significant correlation between feature robustness and prognostic value In

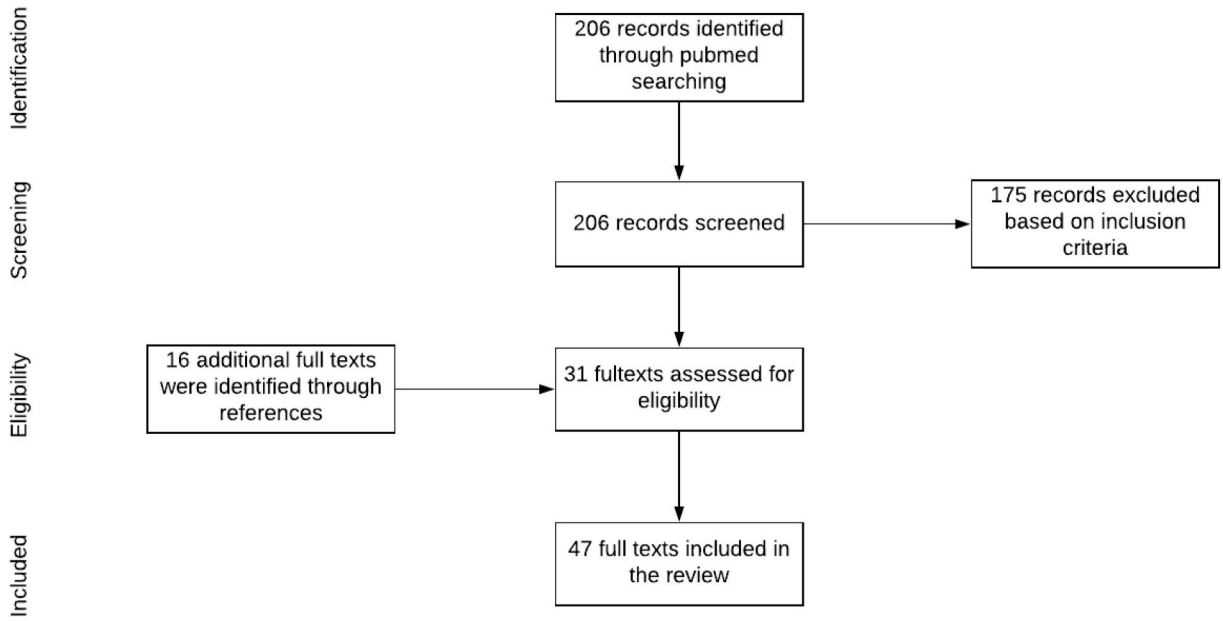

Fig. 1. The PubMed search yielded 206 abstracts for screening against our inclusion criteria, reduced to 31 eligible for full text. The full text was retrieved for 47 abstracts deemed suitable for in-depth evaluation, including 16 that were located in the references of retrieved studies.

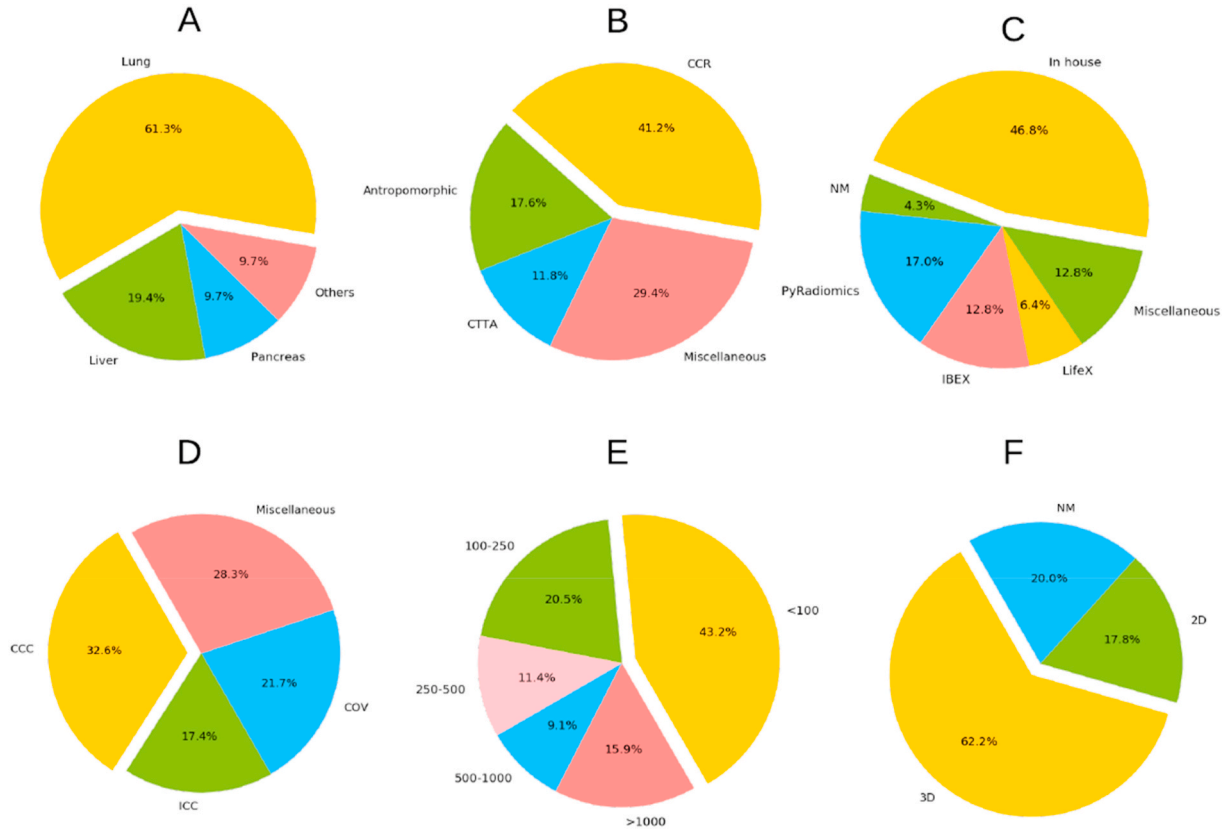

Fig. 2. Reviewed studies based on type of organ studied in human studied (A), common phantom type in phantom studies (B), software used for feature extraction (C), metric used to report features reproducibility (D), distribution of the total number of features extracted (E) and feature dimensions (F).

another study, contrast enhancement in the delayed phase of CT images for NSCLC patients affected some of the radiomic features and the variability of radiomic features due to contrast uptake was found to be dependent largely on the patient characteristics [34].

Reconstruction parameters include reconstruction algorithm (including both reconstruction kernel and algorithm), slice thickness and field of view (including pixel size) were evaluated in 23 studies using data from phantom (5 studies), patient (17 studies) or both (1 studies) and alone or in combination with other parameters (scanner and acquisition) (Supplementary Table 2). Reconstruction algorithm in 10 [7,9,11,24,26,32,44–47] vs. 9 [19,30,31,33,48–52], slice thickness in 17 [5,7,17,22,24,32,33,48–57] vs. none and field of view in 5 [10,22,

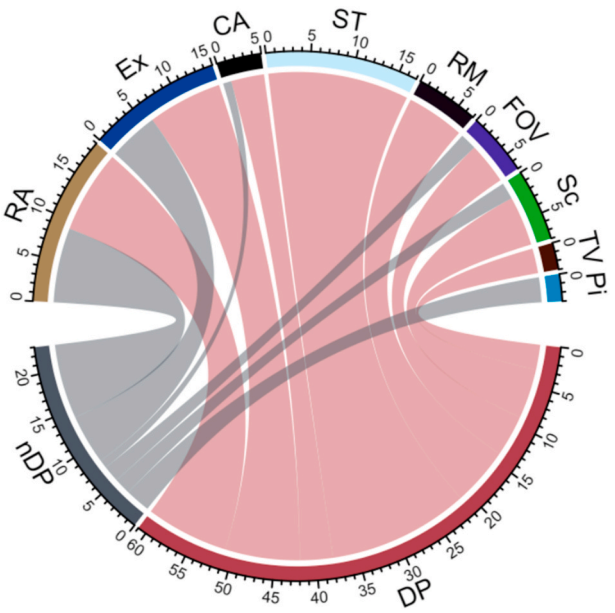

Fig. 3. Distribution of disruptive (DP) and non-disruptive (nDP) parameters reported. Numbers on the scale show the quantity of paper reported as the adjacent parameter as DP or nDP against the other parameters studied. RA: Reconstruction Algorithm, Ex: Exposure, CA: Contrast Agent, ST: Slice Thickness, RM: Respiratory Motion, FOV: Field of View, Sc: Scanner, TV: Tube Voltage and Pi: Pitch.

24,27,54] vs. 2 [16,23] studies were respectively reported as the most and least disruptive parameters for radiomic features (Fig. 3). One study [46] investigated the effect on the noise index resulting from reconstruction algorithms and it was found to be a nDP.

3.2. Inconsistencies

The selected studies suffer from the following variations in their methodology:

Software details (application framework used for feature extraction, programming language, and version) varied significantly across the reviewed studies (Fig. 2C). The majority of studies described their software as in-house software mostly based on MATLAB® (The MathWorks®, Natick, MA) [58], PyRadiomics (Python) [59] and IBEX (MATLAB) [60]. Different software platforms have also shown a significant effect on the statistical variation of Radiomic features [61].

The current literature varied with regards to the optimal metric to use for analysis. The metrics encountered were the CCC in 32.6% of studies, COV in 21.7% and ICC in 17.4% of studies. 28.3% of studies used other metrics such as variability index, interquartile range, proportional difference, absolute difference, or linear mixed-effects models. Some studies reported more than a single metric [24]. Nevertheless, One study reported that there is no statistically significant difference when using either the CCC or the ICC metric [62]. The specific cut off values used to segregate stable from unstable features were not always stated and some studies just reported the percentage of robust features with different cut off values regardless of which cut off value was suitable. Among studies that used the CCC and/or ICC metric, the most common cut off values were 0.9. For COV, cut off values included 1%,5%,10% and 20%. Also, the clustering-based metrics have some limited usage such as the Gaussian mixture model homogeneity, completeness, and the V-measure [15]. However, another study found that the choice of metric as well as the threshold influenced the results [24].

Across the papers analyzed for this study, there was heterogeneity in the total number of features calculated, the type of features, and the feature dimensions. The range of the total number of features was from 5 [25] to 1766 [36] and the majority fell between <100 (~43%) (Fig. 2E). FO, GLCM, RLM and Shape features were the most common feature types used. One ambiguity in the feature calculations was the feature dimension. However, in 62.2% and 17.8% of studies 3D and 2D features (including averaging 2D features over the slices that cover the segmented ROI instead of real 3D calculation) were used respectively (Fig. 2F). Also 20% of studies did not disclose the information regarding feature dimension. A study compared a shape feature and FO features computed from 2D and 3D images and found that the 3D features were more robust than the 2D features across all imaging parameters [7].

4. Discussion

This review investigates the robustness of hand-engineered radiomic features in the context of varying imaging parameters through a deep-dive literature review. In order to make this feasible, imaging parameters, which affect the feature robustness, were collected and grouped into the disruptive and non-disruptive parameters.

Overall, the number of studies addressing the impacts of CT imaging parameters on the robustness of radiomic features is low with the most of these studies being either phantom or retrospective human studies. Although phantom studies are a good guide for conducting prospective studies, the applicability of these studies is limited because of the inherent significant difference between phantom material and human tissue. Even features that we acquired under similar acquisition conditions (such as uniform water phantom) are different from human tissues [31]. Moreover, there is a lack of comprehensive prospective human studies. This is particularly true with respect to investigating the effects of image acquisition parameters. This is due to an inability to rescan patients in the absence of a clinical indication. The drawbacks of the retrospective studies are that the investigators did not have control over the parameters studied, and the range of the scan acquisition parameter variations were limited to those used in imaging patients.

We also aimed to collect a list of robust features along with the most and least disruptive parameters. We reasoned that the resulting list could help future radiomics studies. A comparison study showed that the prognostic CT radiomics model to predict overall survival for locally advanced NSCLC patients based on heterogeneous multicentric dataset with robust features performed equally well as a model trained on a standardized imaging protocol [45]. In another study the hierarchical cluster analysis, which enables the classification of liver lesions with similar radiomic features, showed improved clustering reproducibility when robust radiomic features were used compared with all radiomic features [50]. However, this review is unable to provide such a list due to the substantial inconsistencies related to the reporting style of the reviewed studies. We found that the most common approach to reporting the robust features were the percent of robust features, the robust features against all the imaging parameters and the robust feature-parameter that determine which features are robust against which parameter(s). Nevertheless, Shape [7,24,25,46,51] and FO [11,24,31,35] features were the most reported robust feature classes across the reviewed papers (Supplementary Table 2). We attribute such findings to the fact that features rely on the segmented tumor boundaries, or alternatively, features have low-frequency change components, such as shape. In one study, GLRLM was found to be robust against the reconstruction algorithm [19]. Other studies reported some features individually against specific imaging parameters: GLSZM-SAE (slice thickness) [17], GLRLM-SRE (respiratory motion) [39], GLCM-dissimilarity (slice thickness), FO-Kurtosis, LGRE(exposure), FO-Correlation (reconstruction algorithm) [30], GLRLM-nonuniformity (respiratory motion), GLCM-sumAverage(slice thickness) [5].

The lack of discussion on the clinical implication of feature robustness is another shortfall of the available studies. In a study [52] for predicting epidermal growth factor receptor (EGFR) mutation in lung adenocarcinoma patients, it was concluded that an imaging dataset with uniform imaging parameters is considerably more efficient than

building a model upon robust features extracted from a dataset with non-uniform imaging parameters. Another study showed that models trained and tested on the same slice thickness had better accuracy for the prediction of cirrhotic and healthy liver [5] while the feature stability, expressed by its coefficient of variation, was not considerably influenced by the slice thickness. Moreover, the sensitivity of radiomic features to imaging parameters was shown to be inherently organ dependent [24,35,63,64], patient-specific [34] and even age dependent [65]. It was shown that the number of robust features between normal tissue-derived and tumor-derived features within the same patient is different even with the same imaging parameters [10,27]. These results recommend that studies concerning the robustness of radiomics features should be followed by outcome prediction.

Previous studies have shown that proper pre-processing as suggested by International Biomarker Standard Initiative (IBSI) [66] reduced the impact of image parameter variation on feature robustness [12,17]. BSI is very promising to address the lack of standardized feature definitions, implementation and image pre-processing steps for radiomics by providing reliable benchmark values for commonly used features. However, there is no universal agreement about the type and the details of the preprocessing setting. In a study to evaluate the radiomic feature stability in lung cancer, the researchers found that down-sampling small voxels to large voxels, and thus, creating simple averaging, the process became more desirable compared to up-sampling large voxels to small voxels [18]. They also suggest more precise techniques may be required to ensure optimal pre-processing settings. As shown in previous sections, the robustness of radiomic features depends on the imaging parameters as well. It means global pre-processing settings would not be effective in removing dependencies for all imaging parameters.

Two desirable but less practical solutions of reducing the dependency of radiomic features on imaging parameters are (*i*) credentialing CT scanners used in radiomics studies or correcting for the parameters of the scanner during data analysis; and (ii) adopting standardized image acquisition and reconstruction parameters as is stated by a recent statement of the European Society of Radiology [67]. Recently a statistical harmonization method called ComBat was used in radiomics studies to deal with the impact of imaging parameters [68] by empirical Bayes estimation of prior distribution hyperparameters from standardized features and subsequent estimation of batch effect parameters. However, ComBat fits a model for a specific set of data (standardized features) to transform the data and make them more comparable. Once this is done, the same transformation is not applicable to brand new data from a new imaging center where the correct transformation might be different and must be learned. Providing availability of reliable statistical approach, the most practical way to address the effect of variation in imaging parameters would be using selected reproducible radiomic features [50]. In the meantime, deep learning may allow the use of heterogeneous CT datasets for radiomics studies. In a study CNN model was developed to convert reconstruction kernels and were subsequently able to show improved feature robustness [11]. Another study used a neural network and adversarial training to reduce the disruptive effect of scanner and acquisition parameters on image markers by feature transformation [15].

5. Conclusion

Radiomic features stand to play an important role in guiding personalized cancer treatment. However, the robustness of these features against variation in medical imaging parameters may affect the performance of the models built upon these features. We concluded that, based on the available literature, radiomics features are dependent on imaging parameters. Nevertheless, the impact of this dependency must be evaluated on the prediction of clinical outcome. We have also concluded that there are discrepancies in methodology across the available studies imposing the need for comprehensive studies to further investigate the effects of imaging parameters. We would also suggest establishing a universal methodology and reporting styles in radiomics studies that would make researchers' work more reproducible and create consistency within the scientific community.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Research reported in this paper was supported by the Princess Margaret Cancer Foundation and Canadian Institutes of Health Research (CIHR), Canada under grant number: 426366.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104400.

References

[1] C. Liguori, G. Frauenfelder, C. Massaroni, et al., Emerging clinical applications of computed tomography, Med Devices 8 (2015) 265–278.
[2] Z. Liu, S. Wang, D. Dong, et al., The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges, Theranostics 9 (5) (2019) 1303–1322.
[3] P. Afshar, A. Mohammadi, K.N. Plataniotis, A. Oikonomou, H. Benali, From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities, IEEE Signal Process. Mag. 36 (4) (2019) 132–160.
[4] A. Hosny, C. Parmar, J. Quackenbush, L.H. Schwartz, H.J.W.L. Aerts, Artificial intelligence in radiology, Nat. Rev. Canc. 18 (8) (2018) 500–510.
[5] D. Duda, M. Kretowski, J. Bezy-Wendling, Effect of slice thickness on texture-based classification of liver dynamic CT scans, in: Computer Information Systems and Industrial Management, Springer Berlin Heidelberg, 2013, pp. 96–107.
[6] V. Kumar, Y. Gu, S. Basu, et al., Radiomics: the process and the challenges, Magn. Reson. Imaging 30 (9) (2012) 1234–1248.
[7] B. Zhao, Y. Tan, W.-Y. Tsai, et al., Reproducibility of radiomics for deciphering tumor phenotype with imaging, Sci. Rep. 6 (2016) 23428.
[8] A. Traverso, L. Wee, A. Dekker, R. Gillies, Repeatability and reproducibility of radiomic features: a systematic review, Int. J. Radiat. Oncol. Biol. Phys. 102 (4) (2018) 1143–1158, https://doi.org/10.1016/j.ijrobp.2018.05.053.
[9] J. Solomon, A. Mileto, R.C. Nelson, K. Roy Choudhury, E. Samei, Quantitative features of liver lesions, lung nodules, and renal stones at multi-detector row CT examinations: dependency on radiation dose and reconstruction algorithm, Radiology 279 (1) (2016) 185–194.
[10] R. Yamashita, T. Perrin, J. Chakraborty, et al., Radiomic feature reproducibility in contrast-enhanced CT of the pancreas is affected by variabilities in scan parameters and manual segmentation, Eur. Radiol. (August 7, 2019), https://doi.org/10.1007/s00330-019-06381-8.
[11] J. Choe, S.M. Lee, K.-H. Do, et al., Deep learning-based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses, Radiology 292 (2) (2019) 365–373.
[12] M. Shafiq-Ul-Hassan, K. Latifi, G. Zhang, G. Ullah, R. Gillies, E. Moros, Voxel size and gray level normalization of CT radiomic features in lung cancer, Sci. Rep. 8 (1) (2018) 10545.
[13] J.E. Park, H.S. Kim, D. Kim, et al., A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features, BMC Canc. 20 (1) (2020), https://doi.org/10.1186/s12885-019-6504-5.
[14] D. Mackin, X. Fave, L. Zhang, et al., Measuring computed tomography scanner variability of radiomics features, Invest. Radiol. 50 (11) (2015) 757–765.
[15] V. Andrearczyk, A. Depeursinge, H. Müller, Neural network training for cross-protocol radiomic feature standardization in computed tomography, J. Med. Imaging 6 (2) (2019), 024008.
[16] M. Shafiq-Ul-Hassan, G.G. Zhang, K. Latifi, et al., Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels, Med. Phys. 44 (3) (2017) 1050–1062.
[17] R.T.H.M. Larue, J.E. van Timmeren, E.E.C. de Jong, et al., Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study, Acta Oncol. 56 (11) (2017) 1544–1553.
[18] S.-H. Lee, H.-H. Cho, H.Y. Lee, H. Park, Clinical impact of variability on CT radiomics and suggestions for suitable feature selection: a focus on lung cancer, Canc. Imag. 19 (1) (2019) 54.
[19] M. Kolossváry, B. Szilveszter, J. Karády, Z.D. Drobni, B. Merkely, P. Maurovich-Horvat, Effect of image reconstruction algorithms on volumetric and radiomic

parameters of coronary plaques, J Cardiovasc Comput Tomogr (November 12, 2018), https://doi.org/10.1016/j.jcct.2018.11.004.
[20] D. Moher, L. Shamseer, M. Clarke, et al., Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement, Syst. Rev. 4 (2015) 1.
[21] V. Nardone, A. Reginelli, C. Guida, et al., Delta-radiomics increases multicentre reproducibility: a phantom study, Med. Oncol. 37 (5) (2020) 38.
[22] B.A. Varghese, D. Hwang, S.Y. Cen, et al., Reliability of CT-based texture features: phantom study, J. Appl. Clin. Med. Phys. 20 (8) (2019) 155–163.
[23] M. Biondi, E. Vanzi, G. De Otto, S.F. Carbone, V. Nardone, F. Banci Buonamici, Effects of CT FOV displacement and acquisition parameters variation on texture analysis features, Phys. Med. Biol. 63 (23) (2018) 235021.
[24] R. Berenguer, M.D.R. Pastor-Juan, J. Canales-Vázquez, et al., Radiomics of CT features may Be nonreproducible and redundant: influence of CT acquisition parameters, Radiology 288 (2) (2018) 407–415.
[25] K. Yasaka, H. Akai, D. Mackin, et al., Precision of quantitative computed tomography texture analysis using image filtering: a phantom study for scanner variability, Medicine 96 (21) (2017), e6993.
[26] M. Shafiq-ul-Hassan, G.G. Zhang, D.C. Hunt, et al., Accounting for reconstruction kernel-induced variability in CT radiomic features using noise power spectra, J. Med. Invest. 5 (1) (2017), 011013.
[27] T. Perrin, A. Midya, R. Yamashita, et al., Short-term reproducibility of radiomic features in liver parenchyma and liver malignancies on contrast-enhanced CT imaging, Abdom Radiol (NY) 43 (12) (2018) 3271–3278.
[28] J. Liu, H. Xu, H. Qing, et al., Comparison of radiomic models based on low-dose and standard-dose CT for prediction of adenocarcinomas and benign lesions in solid pulmonary nodules, Front Oncol 10 (2020) 634298.
[29] I. Zhovannik, J. Bussink, A. Traverso, et al., Learning from scanners: bias reduction and feature correction in radiomics, Clin Transl Radiat Oncol 19 (2019) 33–38.
[30] A. Midya, J. Chakraborty, M. Gönen, R.K.G. Do, A.L. Simpson, Influence of CT acquisition and reconstruction parameters on radiomic feature reproducibility, J. Med. Imaging 5 (1) (2018), 011020.
[31] P. Lo, S. Young, H.J. Kim, M.S. Brown, M.F. McNitt-Gray, Variability in CT lung-nodule quantification: effects of dose reduction and reconstruction methods on density and texture based features, Med. Phys. 43 (8) (2016) 4854.
[32] Y.J. Kim, H.-J. Lee, K.G. Kim, S.H. Lee, The effect of CT scan parameters on the measurement of CT radiomic features: a lung nodule phantom study, Comput Math Methods Med 2019 (2019) 8790694.
[33] D.Y. Chong, H.J. Kim, P. Lo, et al., Robustness-driven feature selection in classification of fibrotic interstitial lung disease patterns in computed tomography using 3D texture features, IEEE Trans. Med. Imag. 35 (1) (2016) 144–157.
[34] R. Kakino, M. Nakamura, T. Mitsuyoshi, et al., Comparison of radiomic features in diagnostic CT images with and without contrast enhancement in the delayed phase for NSCLC patients, Phys. Med. 69 (2020) 176–182.
[35] D. Mackin, R. Ger, C. Dodge, et al., Effect of tube current on computed tomography radiomic features, Sci. Rep. 8 (1) (2018) 2354.
[36] X. Li, E. Chen, B. Guo, et al., The impact of respiratory motion and CT pitch on the robustness of radiomics feature extraction in 4DCT lung imaging, Comput. Methods Progr. Biomed. 197 (2020) 105719.
[37] Q. Du, M. Baine, K. Bavitz, et al., Radiomic feature stability across 4D respiratory phases and its impact on lung tumor prognosis prediction, PloS One 14 (5) (2019), e0216480.
[38] L.A. Hunter, S. Krafft, F. Stingo, et al., High quality machine-robust image features: identification in nonsmall cell lung cancer computed tomography images, Med. Phys. 40 (12) (2013) 121916, https://doi.org/10.1118/1.4829514.
[39] K. Lafata, J. Cai, C. Wang, J. Hong, C.R. Kelsey, F.-F. Yin, Spatial-temporal variability of radiomic features and its effect on the classification of lung cancer histology, Phys. Med. Biol. 63 (22) (2018) 225003.
[40] E. Huynh, T.P. Coroller, V. Narayan, et al., Associations of radiomic data extracted from static and respiratory-gated CT scans with disease recurrence in lung cancer patients treated with SBRT, PloS One 12 (1) (2017), e0169172.
[41] R.T.H.M. Larue, L. Van De Voorde, J.E. van Timmeren, et al., 4DCT imaging to assess radiomics feature stability: an investigation for thoracic cancers, Radiother. Oncol. 125 (1) (2017) 147–153.
[42] I.S. Gruzdev, K.A. Zamyatina, V.S. Tikhonova, et al., Reproducibility of CT texture features of pancreatic neuroendocrine neoplasms, Eur. J. Radiol. 133 (2020) 109371.
[43] Y.-H. Zhang, K. Brehmer, A. Svensson, G. Herlin, P. Stål, T.B. Brismar, Variation in textural parameters of hepatic lesions during contrast medium injection, Acta Radiol. (October 27, 2020), 284185120964904.
[44] S. Denzler, D. Vuong, M. Bogowicz, et al., Impact of CT convolution kernel on robustness of radiomic features for different lung diseases and tissue types, Br. J. Radiol. (February 5, 2021), 20200947.
[45] D. Vuong, M. Bogowicz, S. Denzler, et al., Comparison of robust to standardized CT radiomics models to predict overall survival for non-small cell lung cancer patients, Med. Phys. 47 (9) (2020) 4045–4053.
[46] H. Kim, C.M. Park, M. Lee, et al., Impact of reconstruction algorithms on CT radiomic features of pulmonary tumors: analysis of intra- and inter-reader variability and inter-reconstruction algorithm variability, PloS One 11 (10) (2016), e0164924.
[47] J.J. Foy, M. Shenouda, S. Ramahi, S. Armato, D.T. Ginat, Effect of an iterative reconstruction quantum noise reduction technique on computed tomography radiomic features, J. Med. Imaging 7 (6) (2020), 064007.
[48] B.S. Erdal, M. Demirer, K.J. Little, et al., Are quantitative features of lung nodules reproducible at different CT acquisition and reconstruction parameters? PloS One 15 (10) (2020), e0240184.
[49] F. Prayer, J. Hofmanninger, M. Weber, et al., Variability of computed tomography radiomics features of fibrosing interstitial lung disease: a test-retest study, Methods (September 4, 2020), https://doi.org/10.1016/j.ymeth.2020.08.007.
[50] M. Meyer, J. Ronald, F. Vernuccio, et al., Reproducibility of CT radiomic features within the same patient: influence of radiation dose and CT reconstruction settings, Radiology 293 (3) (2019) 583–591.
[51] L. Lu, R.C. Ehmke, L.H. Schwartz, B. Zhao, Assessing agreement between radiomic features computed for multiple CT imaging settings, PloS One 11 (12) (2016), e0166550.
[52] Y. Li, L. Lu, M. Xiao, et al., CT slice thickness and convolution kernel affect performance of a radiomic model for predicting EGFR status in non-small cell lung cancer: a preliminary study, Sci. Rep. 8 (1) (2018) 17913.
[53] B.A. Varghese, D. Hwang, S.Y. Cen, et al., Identification of robust and reproducible CT-texture metrics using a customized 3D-printed texture phantom, J. Appl. Clin. Med. Phys. 22 (2) (2021) 98–107.
[54] S. Loi, M. Mori, G. Benedetti, et al., Robustness of CT radiomic features against image discretization and interpolation in characterizing pancreatic neuroendocrine neoplasms, Phys. Med. 76 (2020) 125–133.
[55] H.-T. Hu, Q.-Y. Shan, S.-L. Chen, et al., CT-based radiomics for preoperative prediction of early recurrent hepatocellular carcinoma: technical reproducibility of acquisition and scanners, Radiol. Med. 125 (8) (2020) 697–705.
[56] S. Park, S.M. Lee, K.H. Do, et al., Deep learning algorithm for reducing CT slice thickness: effect on reproducibility of radiomic features in lung cancer, Korean J. Radiol. 20 (10) (2019) 1431–1440.
[57] P. Guggenbuhl, D. Chappard, M. Garreau, J.-Y. Bansard, G. Chales, Y. Rolland, Reproducibility of CT-based bone texture parameters of cancellous calf bone samples: influence of slice thickness, Eur. J. Radiol. 67 (3) (2008) 514–520.
[58] MathWorks - makers of MATLAB and simulink. https://www.mathworks.com/. (Accessed 17 May 2020).
[59] van Griethuysen JJM, A. Fedorov, C. Parmar, et al., Computational radiomics system to decode the radiographic phenotype, Canc. Res. 77 (21) (2017) e104–e107.
[60] L. Zhang, D.V. Fried, X.J. Fave, L.A. Hunter, J. Yang, L.E. Court, IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics, Med. Phys. 42 (3) (2015) 1341–1353.
[61] I. Fornacon-Wood, H. Mistry, C.J. Ackermann, et al., Reliability and Prognostic Value of Radiomic Features Are Highly Dependent on Choice of Feature Extraction Platform, European Radiology, 2020, https://doi.org/10.1007/s00330-020-06957-9.
[62] P. Hu, J. Wang, H. Zhong, et al., Reproducibility with repeat CT in radiomics study for rectal cancer, Oncotarget 7 (44) (2016), https://doi.org/10.18632/oncotarget.12199.
[63] R.N. Mahon, G.D. Hugo, E. Weiss, Repeatability of texture features derived from magnetic resonance and computed tomography imaging and use in predictive models for non-small cell lung cancer outcome, Phys. Med. Biol. (April 12, 2019), https://doi.org/10.1088/1361-6560/ab18d3.
[64] X. Fave, D. Mackin, J. Yang, et al., Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? Med. Phys. 42 (12) (2015) 6784–6797.
[65] S. Boughdad, C. Nioche, F. Orlhac, L. Jehl, L. Champion, I. Buvat, Influence of age on radiomic features in 18F-FDG PET in normal breast tissue and in breast cancer tumors, Oncotarget 9 (56) (2018) 30855–30868.
[66] A. Zwanenburg, S. Leger, M. Vallières, S. Löck, Image Biomarker Standardisation Initiative; 2016, ArXiv161207003 Cs, 2019.
[67] (esr) ES of R, European Society of Radiology (ESR), ESR statement on the validation of imaging biomarkers, Insights into Imaging 11 (1) (2020), https://doi.org/10.1186/s13244-020-00872-9.
[68] R.N. Mahon, M. Ghita, G.D. Hugo, E. Weiss, ComBat harmonization for radiomic features in independent phantom and lung cancer patient computed tomography datasets, Phys. Med. Biol. 65 (1) (2020), 015010.