Evaluation of the Use of Radiomics in Predicting Treatment Response in Oncology:

(Review Article)

NASSER SIHLI ALSHAMMARY, MOHAMMED NUHAYR ALWAHDANI, FARES MOTALQ ALONAZI, FAHAD MADALLAH ALNUWMASIU and SHAYEM HAMDAN ALHARBI

KSA, National Guard Health Affairs

Abstract

Background: Radiomics involves the retrieval of numerical information from medical imaging, with the ability to describe the characteristics of a tumor. The radiomics technique has the ability to create prognostic models for therapy response, which is crucial for the advancement of personalized medicine.

Aim of Study: This literature review provides a concise overview and assesses the scientific rigor and reporting standards of radiomics research in predicting treatment response in non-small-cell lung cancer (NSCLC).

Methods: An extensive literature search was performed using the PubMed database. The radiomics quality score (RQS), a measure specifically designed for radiomics, was used to evaluate the scientific and reporting quality, following the parameters set by TRIPOD.

Results: The studies included in the analysis revealed several predictive markers, including first-, second-, and high-order features. These characteristics included kurtosis, grey-level uniformity, and wavelet HLL mean, as well as PET-based metabolic indicators. The studies exhibited significant variability as a result of variations in patient demographics, cancer stage, treatment methods, duration of follow-up, and radiomics processing protocols.

Conclusion: The use of radiomics research in clinical practice has not yet been implemented. To develop radiomic predictors of response that can be reproduced, it is necessary to make efforts toward standardization and cooperation. In order for radiomic models to be used as a clinical decision-making tool for personalized treatment of patients with NSCLC, it is necessary to verify them externally and assess their effect within the therapeutic pathway.

Key Words: Radiomics – Review – Non-Small-Cell Lung Cancer (NSCLC) – Radiomics Quality Score – Treatment.

Introduction

RADIOMICS refers to the process of extracting data from medical imaging via the use of mathematical algorithms for the purpose of conducting sophisticated picture analysis [1]. Radiomics is based on the idea that medical imaging captures quantitative data that cannot be seen by the human eye, but that may indicate the underlying tissue’s pathology. Quantitative radiomic features in cancer imaging have the ability to accurately describe the characteristics of a tumor’s phenotype. The primary objective of radiomics is to develop prognostic models for therapy response by analyzing the tumor phenotypic features obtained from medical imaging. It is crucial for the advancement of personalized medicine, which involves customizing therapy based on the specific attributes of particular patients and their tumors.

Lung cancer is the prevailing form of cancer globally and the primary cause of cancer-related mortality. According to data from 2018, there were 2.09 million cases of lung cancer diagnosed and 1.76 million deaths caused by lung cancer [2]. Non-small-cell lung carcinoma (NSCLC) is the predominant form of lung cancer, representing 87% of all diagnosed cases [3]. Various therapeutic methods are used in the management of NSCLC, including surgery, radiation (including stereotactic ablative radiotherapy), and systemic therapy (such as cytotoxic chemotherapy, tyrosine kinase inhibitors, and immune checkpoint inhibitors) [4]. Patients diagnosed with non-small cell lung cancer (NSCLC) undergo initial diagnostic and staging imaging using computed tomography (CT) and/or fluorodeoxyglucose positron emission tomography/computed tomography (FDG PET/CT). Periodic imaging is also conducted to assess the effectiveness of therapy and check for any reappearance.
Either pathologic or radiologic criteria may be used to evaluate the effectiveness of therapy. Pathologic response is a definitive outcome, but it can only be assessed in the minority of patients (16%) with non-small cell lung cancer (NSCLC) who have surgery to remove the tumor [5]. Assessing the response to therapy in NSCLC mostly relies on evaluating radiologic response. RECIST, which stands for Response Evaluation Criteria in Solid Tumours, offers a precise and standardized approach to assessing the effectiveness of treatment by measuring the size of the tumor in a one-dimensional manner [6]. The RECIST criteria are included into the establishment of oncology trial endpoints, such as response rate and progression-free survival [7]. In practical practice, the assessment of treatment response using radiology mostly depends on the size of the tumor, along with a qualitative evaluation of other tumor features such as homogeneity and form.

From a quantitative perspective, this technique is not only rudimentary but also disregards a significant amount of information included in the medical imaging. The radiomics technique has the capacity to detect quantitative indicators of therapeutic response at an earlier stage of treatment. This may facilitate the adjustment, intensification, or modification of therapy at an earlier stage of the illness in order to enhance patient outcomes.

**Aim of Work:**

Despite its potential, radiomics has not yet been used in clinical practice [8]. This literature review provides a concise overview of the current uses of radiomics in predicting therapy response in non-small cell lung cancer (NSCLC). Additionally, it assesses the scientific rigor and reporting standards of research conducted in this area. Prior studies in this domain have mostly examined innovative radiomic procedures [9] and forecasting prognosis [10]. However, our primary objective is to predict therapeutic response at an earlier stage of therapy. To the best of our knowledge, this is the first study to assess the research quality in this particular subject. We analyze the research obstacles that contribute to the translational gap in radiomics and speculate about potential future paths.

**Methods**

A thorough literature search was performed using the PubMed database, including a diverse set of keywords and Medical Subject Headings (MeSH) phrases. We included all studies that assessed quantitative characteristics derived from initial or early treatment CT or PET/CT scans in relation to treatment response in patients receiving any kind of therapy for NSCLC at any point. The exclusion criteria we used were as follows: Studies that did not evaluate radiologic or pathologic response as an endpoint; studies that solely focused on methodological aspects of radiomics; studies that extracted quantitative features from imaging conducted after treatment, which were not predictive; studies conducted on phantom or animal models; articles that did not contain original data, such as reviews and editorials. Exclusion criteria did not include language, geographical area, or date of publication.

**Anticipating the occurrence of a pathological reaction:**

Pathologic full response is characterized by the absence of malignant cells in all samples. It is a significant predictive factor in locally advanced non-small cell lung cancer (NSCLC) and is linked to improve overall survival and reduced chances of both local and distant recurrence [27]. Three retrospective studies examined the ability of CT-based radiomic characteristics to predict pathologic response. Among patients with non-small cell lung cancer (NSCLC) who had a combination of chemotherapy and radiation followed by surgical removal of the tumor, the wavelet HLL mean, which is a high-order textural feature, showed modest predictive ability for achieving a full response based on pathological examination (area under the curve [AUC] 0.63, p=0.01) [14].

Subsequent research conducted by the same group shown that the texture characteristics of lymph nodes were more effective in predicting pathologic full response compared to the texture features of the main tumor [13]. A predictive model, constructed using ten radiomic features from primary tumors and ten radiomic features from lymph nodes, demonstrated significantly improved accuracy in predicting pathologic response compared to conventional features (AUC 0.68, p<0.05). Additionally, a model that combined clinical and radiomic features performed the best in predicting the presence of gross residual disease (AUC 0.73, p<0.05) [13].

In their study, Chong et al., [15] conducted a multivariate analysis on two groups of patients. One group received combination chemoradiotherapy, while the other received tyrosine kinase inhibitor therapy. Both groups then had surgical resection. The study found that the likelihood of a pathologic response was predicted by kurtosis in patients who received combination chemoradiotherapy (odds ratio 1.107, p=0.009), and by intensity variability in patients who received tyrosine kinase inhibitor therapy (odds ratio 1.093, p=0.028) [15].

Aukema et al., [16] conducted a prospective research to examine the correlation between PET-based quantitative characteristics and pathological response in patients with non-small cell lung cancer (NSCLC) who had combination chemoradiotherapy followed by surgical resection. An early shift in the maximum standardised uptake value (SUVmax) was shown to be a very accurate predictor of pathologic complete response, with a k-agreement of 0.55 and a p-value of 0.008 [16].
Anticipating the radiological outcomes:

The RECIST criteria are often used in cancer studies to assess therapy response by radiologic examination. Recent research including 23,259 cancer patients (with 36% having lung cancer) who received chemotherapy and/or targeted treatments revealed a direct correlation between the change in the size of the tumor in one dimension and the overall survival rate [28]. Assessing changes in tumor volume is a useful method for assessing the effectiveness of radiologic therapy. This method has been shown to be more closely related to pathologic complete response than the unidimensional RECIST method in patients with locally advanced NSCLC [29].

Four studies examined PET-based metabolic parameters. In patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) who were treated with tyrosine kinase inhibitors, the initial PET metabolic parameters, including maximum standardized uptake value (SUVmax) and total lesion glycolysis (TLG), did not show any correlation with the response according to the Response Evaluation Criteria in Solid Tumors (RECIST) [18,19]. Nevertheless, there was a significant correlation between the change in SUVmax during PET imaging from the first assessment to 6 weeks into the therapy and the RECIST response at 12 weeks (p=0.007) [18]. Comparable outcomes were seen in individuals who had simultaneous chemoradiotherapy.

The baseline SUVmax in this group was not a reliable indicator of the RECIST response, as shown by an area under the curve (AUC) of 0.64 [22]. A research conducted on prospective participants showed that the change in maximum standardized uptake value (SUVmax) and the net-influx constant (Ki) of fluorodeoxyglucose (FDG) during serial positron emission tomography (PET) imaging, from the first scan to the first cycle of chemotherapy, was very accurate in predicting the response according to the Response Evaluation Criteria in Solid Tumors (RECIST) guidelines. The area under the curve (AUC) for SUVmax and Ki were 0.91 and 0.92, respectively [23].

Three studies examined characteristics related to texture and heterogeneity using PET imaging. In patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) who received tyrosine kinase inhibitor treatment, some characteristics indicating tumor heterogeneity seen during baseline PET/CT imaging, such as first-order standard deviation, entropy, and uniformity, were shown to be significantly linked with RECIST response (p<0.01) [18]. Among patients undergoing concurrent chemoradiotherapy, certain textural variables at baseline, including contrast, coarseness, and busyness, was shown to be predictive of RECIST response. The area under the curve (AUC) values for these features were 0.80, 0.82, and 0.72, respectively, with a p-value of less than 0.03 [20].

A separate study demonstrated that the contrast and coefficient of variation of SUV at the beginning of the study (AUC 0.80 and 0.78 respectively), as well as the change in contrast and coefficient of variation of SUV over consecutive PET scans from the start of the study to 4 weeks into treatment (AUC 0.86 and 0.80 respectively), were both indicative of the RECIST response at 12 weeks [17].

In patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) who received first-line platinum-based chemotherapy, the uniformity of grey levels on baseline CT scans was shown to be predictive of RECIST response. This prediction was particularly accurate in a subset of patients with adenocarcinoma histology, with an area under the curve (AUC) of 0.741 and a p-value of less than 0.01 [21]. Ramella et al. [24] demonstrated that the combination of seven radiomic variables taken from pre-treatment CT scans and five conventional clinical features accurately predicted tumor volume following concurrent chemoradiotherapy, with an area under the curve (AUC) of 0.82 [24].

Other organizations have constructed prediction models in order to facilitate adaptive radiotherapy. A predictive model, based on 35 quantitative parameters derived from CT scans taken before treatment, showed a strong correlation (r 0.83) with tumor volume after 6 weeks of radiation [26]. In their study, Zhang et al. [25] demonstrated that a model using pre-treatment CT characteristics together with mid-treatment CT for “mid-course correction” was an effective predictor of post-radiotherapy tumor volume. This prediction model was further verified in an independent cohort and achieved an area under the curve (AUC) of 0.85 [25].

Discussion

This study provides a summary of fourteen studies that examine the use of CT and PET/CT radiomic predictors to determine therapy response in NSCLC. An assessment has been conducted to evaluate the scientific rigor and reporting standards of these researches. The studies included in the analysis reported a range of predictive markers, including histogram-based properties like kurtosis [15], second-order textural features like grey-level uniformity [21], high-order features like wavelet HLL mean [14], and features that describe changes in the PET-based net-influx rate constant (Ki) [23].

Only a few studies found the same radiomic characteristic to be predictive of treatment response in NSCLC. This may be partially attributed to the significant variation across different research. The investigations were conducted in several patient groups with varying cancer stages, being treated at different institutions, and with variable timeframes for follow-up imaging. Differences in image capture and reconstruction procedures across institu-
tions may lead to variations in quantitative imaging characteristics that are unrelated to biological factors. The research examined the effect of various treatment methods on disease response, such as traditional radiation, cytotoxic chemotherapy, simultaneous chemoradiotherapy, and tyrosine kinase inhibitors. Considering their distinct methods of action, it is physiologically feasible that the radiomic indicators of response vary for each modality.

In their 2014 study, Chong et colleagues examined patients with stage IIIA NSCLC at a single institution. They revealed distinct radiomic indicators of pathological response in patients treated with chemotherapy and those treated with tyrosine kinase inhibitors [15]. As far as we know, there has been no study on the radiomic predictors of pathologic or radiologic response in NSCLC patients who have had immunotherapy.

The study quality varied across the fourteen included studies, as shown by their scores on the RQS metric ranging from −5 to +9. The highest attainable score on this metric was +36. The studies included in the analysis had a very poor level of scientific rigor and reporting quality. This may have led to an increased chance of reporting a false-positive correlation between radiomic characteristics and treatment response.

Out of all the studies conducted, only three were prospective. These three studies had the advantage of having standardized cancer stage, therapy, and follow-up procedures [16,22,23]. Although most research provided detailed information on their imaging techniques, only a small number of studies included test-retest imaging, phantom imaging, and repeated segmentation to evaluate the reliability of the features. Only characteristics with a high level of consistency and accuracy can effectively demonstrate the fundamental biological properties of tissue and so serve as indicators of how a therapy will be effective [24-26]. Two studies included in the analysis found that coarseness and contrast are strong predictors of RECIST response, with area under the curve (AUC) values of 0.80 or higher [17,20]. However, it is worth noting that these two higher order characteristics have been shown to be among the least repeatable radiomic variables [27-30]. These variables increase the likelihood of bias in the positive outcomes reported by individual research, making them less likely to be applicable to a wider population.

In several researches, appropriate techniques for reducing characteristics were used. For instance, Coroller et al., [13,14] eliminated highly correlated and non-reproducible data before doing the analysis. In the absence of proper feature reduction, several studies were prone to overfitting. For instance, Hunter et al., [26] developed a model employing 35 radiomic characteristics and a sample size of 64 patients [26]. It is well acknowledged that in order for a model to be generalizable, a minimum of 10 patients per radiomic characteristic is necessary [1]. The studies mostly focused on evaluating the effectiveness of radiomic markers using discrimination statistics, while neglecting the use of calibration data. Out of all the studies that reported cutoff analyses, only two of them employed a pre-determined threshold [16,23]. Using a post hoc optimum cutoff selection method together with a high number of potential radiomic characteristics has been shown to greatly raise the likelihood of type I error (with a 76% chance) [31,32]. All studies save for one [28] lacked external validation. Collectively, these characteristics have probably resulted in too optimistic predictions in several research that were included.

Attempts were attempted to establish a connection between radiomic properties and biological traits; nevertheless, there was a significant absence of information about the practical use in clinical settings. No research conducted a clinical utility analysis or cost-effectiveness analysis. Only two studies [13,14] included an assessment of the additional benefit provided by radiomics compared to the existing ‘gold standard’. Thus, it is not unexpected that the suggested radiomic predictors have not been used in clinical practice.

This review is commendable for its comprehensive coverage of the literature, concise presentation of research findings, and use of a standardized quality rating technique. A significant constraint is that this is not a comprehensive assessment conducted according to a certain methodology. It is probable that several pertinent research have been omitted due to the lack of searching in other medical databases and the grey literature. Due to the heterogeneity across individual studies, it was not possible to do a meta-analysis.

Some claim that the conventional radiomics technique is less effective compared to artificial intelligence (AI) technologies like deep learning using convolutional neural networks [31]. This is because AI eliminates the need for manual feature extraction and selection, which might introduce human bias. Nevertheless, AI methods need much bigger sets of annotated imaging records, and the rationale behind the judgments made by the AI system remains opaque or not fully understood [33]. The integration of conventional radiomics with artificial intelligence (AI) has the potential to use the strengths of both approaches [34]. Radiologists play a vital role in organizing high-quality imaging information by using a consistent and organized reporting vocabulary. This helps to conduct extensive investigations involving large groups of individuals. Indeed, this poses a significant challenge in the context of everyday therapeutic practice.
References

1- GILLIES R.J., KINAHAN P.E. and HRICAK H.: Radiomics: Images are more than pictures, they are data. Radiology, 278: 563-577, 2016.


Evaluation of the Use of Radiomics in Predicting Treatment Response in Oncology


تقييم استخدام تقنية الراديومكس في توقع استجابة العلاج

تغطي تقنية الراديومكس استرجاع المعلومات العددية من الصور الطبية، مع القدرة على وصف خصائص الورم. تتمتع تقنية الراديومكس بالقدرة على إنشاء نماذج توقعية لأستجابة العلاج، وهو أمر حاسم لتقديم الطب الشخصي.

هدف العمل: يوفر هذا الاستعراض الأدبي نظرة عامة موجزة وзначمة الدقة العلمية ومعايير التقارير في البحوث المتعلقة بتقنية الراديومكس في توقع استجابة العلاج في سرطان الرئة غير الصغير الخلايا (NSCLC).


النتائج: كشفت الدراسات الشاملة في التحليل عن عدة علامات توقعية، بما في ذلك المؤشرات من الدرجة الأولى والثانية، وال عالية. تشمل هذه المواد الأعراض الفيزيولوجية، وراء مستوى الوقود الرمادي، ومتوسط موجة Hzz بالإضافة إلى المؤشرات الاستجابة العظمى على الصورة بالتصوير بالاستخدام الإشعاعي المتغير. وقد أظهرت الدراسات تباينًا كبيرًا نتيجة اختلافات في خصائص الفرد والمرحلة السرطان والطرق العلاجية ومدة التالية، وبروتوكولات معالجة الراديومكس.

الاستنتاج: لم يتم تطبيق استخدام البحوث في مجال الراديومكس في الممارسة السريرية. بعد تطوير متوقعات الراديومكس في الدراسة السريرية بعد. تطور متوقعات الراديومكس للاستجابة التي يمكن تكرارها، من الضروري بناء جهد للمتعدد والتعاون. من أجل استخدام نماذج الراديومكس كأداة لاتخاذ القرارات السريرية للعلاج الشخصي المرضى المصابين بـ NSCLC وتحديد مستقبلها خارجياً وتقييم تأثيرها ضمن المسار العلاجي.