



Disponible en ligne sur

ScienceDirect
www.sciencedirect.com

Elsevier Masson France

EM|consulte
www.em-consulte.com



MISE AU POINT

Radiomique : mode d'emploi. Méthodologie et exemples d'application en imagerie de la femme



Radiomics: instructions for use. Methodology and examples of applications in women's imaging

Saskia Vande Perre^{a,*,b,c}, Loic Duron^{b,c,d},
Audrey Milon^a, Stéphanie Nougaret^{e,f,g},
Laure Fournier^{b,c,h}, Isabelle Thomassin-Naggara^{a,i}

^a Département de radiologie, AP-HP hôpital Tenon, 4, rue de la Chine, 75020 Paris, France

^b Université Paris Descartes Sorbonne Paris Cité, 4, rue de la Chine, 75020 Paris, France

^c Inserm, UMR-S970, Cardiovascular Research Center PARCC, 4, rue de la Chine, 75020 Paris, France

^d Département de radiologie, fondation ophtalmologique Adolphe-de-Rothschild, Paris, France

^e Département de radiologie, Institut de recherche en cancérologie de Montpellier (IRCM), campus Val d'Aurelle, 34298 Montpellier cedex 5, France

^f Inserm, U1194, campus Val d'Aurelle, 34298 Montpellier cedex 5, France

^g Université de Montpellier, 34295 Montpellier, France

^h Département de radiologie, AP-HP hôpital européen Georges-Pompidou, 20, rue Leblanc, 75015 Paris France

ⁱ Équipe médecine, Sorbonne Université, université Pierre-et-Marie-Curie Paris 6, ISCD, IUC, 75005 Paris, France

Reçu le 14 janvier 2019 ; accepté le 18 janvier 2019

Disponible sur Internet le 8 mars 2019

MOTS CLÉS

Radiomique ;
Texture ;
Cancer ;
Biomarqueurs ;

Résumé L'étude de paramètres quantitatifs extraits des images radiologiques grâce à des descripteurs mathématiques corrélés à la réalité biologique permettrait la détermination de la nature histologique ou biologique des tissus. Une méthodologie commune aux études de radiomique est nécessaire afin de permettre la validation puis l'application clinique de ces nouveaux biomarqueurs tumoraux. Sept étapes essentielles composent une étude de radiomique : l'acquisition, la segmentation, le prétraitement, l'extraction des données, la

* Auteur correspondant.

Adresse e-mail : s.vandeperre@gmail.com (S. Vande Perre).

Méthodologie ;
Paramètres
quantitatifs

KEYWORDS

Radiomics;
Texture;
Cancer;
Biomarkers;
Methodology;
Quantitative imaging
features

sélection des paramètres, la classification et enfin la validation. L'application de ces méthodes en imagerie de la femme, d'abord en IRM mammaire et plus récemment dans l'étude des cancers pelviens, est prometteuse et pourrait, en apportant de nouveaux biomarqueurs diagnostiques, thérapeutiques et pronostiques, modifier notre pratique dans les années à venir.

© 2019 Elsevier Masson SAS. Tous droits réservés.

Summary Quantitative imaging features cannot be accurately appreciated by the human eye whereas computerized tools may overcome this issue and help discovering new imaging biomarkers. Correlated with the biological reality, these new biomarkers could allow prediction of diagnosis, treatment response and prognosis. A common methodology for radiomics studies is needed to enable validation and clinical application. This article describes the seven steps needed to build a radiomics study: acquisition, segmentation, pre-processing, data extraction, features selection, classification and validation. In women's imaging, two topics were explored: first in breast MRI and more recently in the study of pelvic cancers. The results are promising and could possibly modify our practice in the coming years by bringing new diagnostic, therapeutic and prognostic biomarkers.

© 2019 Elsevier Masson SAS. All rights reserved.

Introduction

La mesure quantitative de paramètres de texture extraits d'images radiologiques caractérisant l'hétérogénéité tumorale à un niveau non visible à l'œil nu et corrélés à la nature biologique des tissus semble émerger comme un nouveau biomarqueur tumoral au même titre que l'hétérogénéité histologique ou moléculaire. Son utilisation pourrait permettre de prédire la nature biologique d'un tissu, la réponse thérapeutique ou le pronostic d'une lésion tumorale.

Le mot de « radiomique », calqué sur la génomique, est apparu pour la première fois dans la littérature en 2012 dans un article publié par Lambin et al. [1]. Il s'agissait d'une méthode d'étude scientifique basée sur les données (data-driven) dont le but était d'extraire un grand nombre de paramètres, parmi lesquels les plus intéressants statistiquement étaient sélectionnés pour ensuite en déduire une hypothèse scientifique. Cette approche est différente de l'approche scientifique classique basée sur l'hypothèse (hypothesis-driven) et qui tend à confirmer ou infirmer une hypothèse préalablement établie [2]. Comme toutes les techniques à haut débit (les « -omics »), cette recherche cherche à optimiser le processus de découverte de nouveaux biomarqueurs en imagerie.

Avec plus de 400 articles publiés depuis 2012, la radiomique est devenue un centre d'intérêt majeur de la recherche en imagerie. Cependant, il s'agit d'une technique très nouvelle qui doit encore être standardisée afin d'assurer sa reproductibilité et sa crédibilité. En effet, de nombreuses équipes ne détaillent pas dans leurs articles les méthodes de prétraitement de l'image, la manière de coder les paramètres quantitatifs ou les étapes de sélection des paramètres qui peuvent être le fruit de choix arbitraires. Un collectif international de chercheurs a récemment publié un guide apportant des définitions et des recommandations de bonnes pratiques dans la recherche et la sélection de biomarqueurs issus des images radiologiques [3], que nous allons détailler dans cet exposé.

Méthodologie

La radiomique nécessite sept étapes [4] : l'acquisition, la segmentation, le prétraitement, l'extraction des paramètres, la sélection des données, la classification et la validation (Fig. 1).

Acquisition

De l'échographie à l'IRM en passant par la mammographie, la radiologie conventionnelle, le scanner ou la tomographie par émission de positons (TEP)-tomodensitométrie (TDM), les images issues de toutes modalités sont utilisables pour l'analyse radiomique [5–8]. Après acquisition, il existe souvent un post-traitement des images à partir des données brutes. Certains post-traitements très spécifiques, afin de réduire le bruit en IRM ou les artefacts métalliques en scanner, devront être renseignés dans l'article autant que possible car ils peuvent impacter les paramètres calculés [9,10]. Ils sont cependant rarement détaillés par les constructeurs, et peuvent varier beaucoup d'une machine ou d'un centre à l'autre pour les données multimachines ou multicentriques.

Segmentation

Après acquisition des images, l'étape suivante consiste à définir dans quelle région l'extraction de données et notamment l'analyse de texture va être réalisée. Différentes méthodes de segmentations sont utilisables. La segmentation manuelle par délinéation, par exemple en utilisant un logiciel comme ITK-SNAP, permet de se rapprocher de la pratique radiologique quotidienne (Fig. 2) [11]. Cette méthode nécessite néanmoins un temps humain important qui limite la constitution des grandes bases de données nécessaires à l'utilisation et la validation de la radiomique. La segmentation automatique fait appel à des logiciels qui détectent automatiquement les contours des lésions, et sont particulièrement performants dans le poumon et le foie.

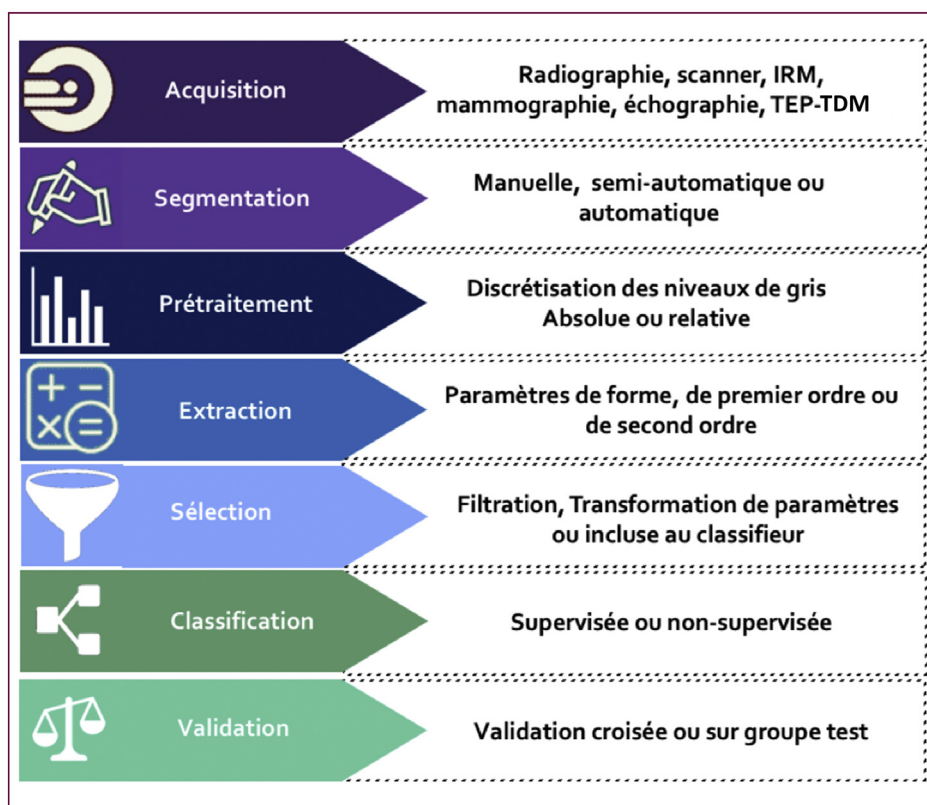


Figure 1. Étapes de la radiomique.

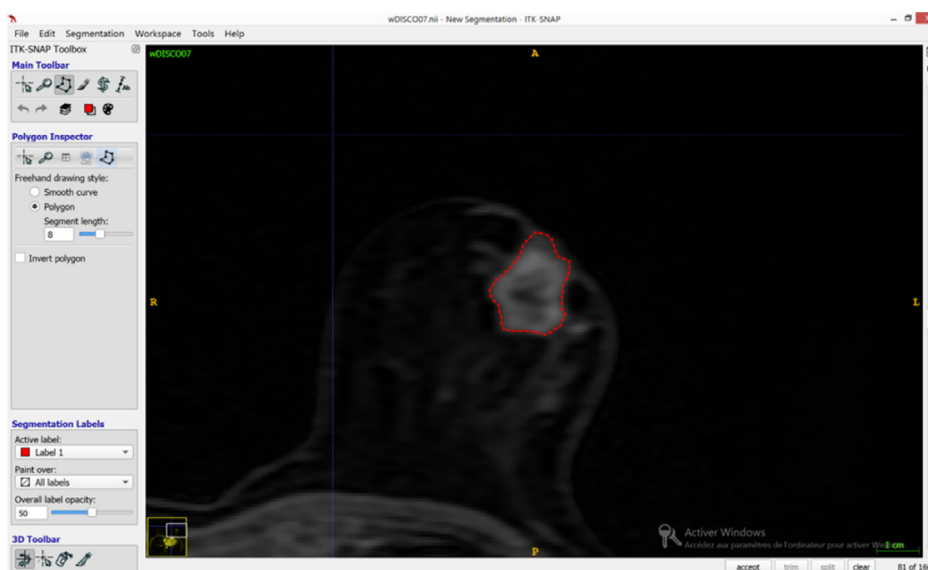


Figure 2. Capture d'écran de la segmentation manuelle d'une tumeur mammaire réalisée grâce au logiciel ITK-SNAP (version 3.6) sur une séquence d'IRM ultrafast après injection de gadolinium.

La segmentation semi-automatique est un intermédiaire entre les deux, où un logiciel définit une délimitation, qui sera ensuite ajustée par l'utilisateur [12,13]. La méthode de segmentation a un impact sur les résultats des paramètres de texture. Parmar et al. ont ainsi montré qu'une segmentation semi-automatique permettait une meilleure reproductibilité des paramètres de texture qu'une segmentation manuelle [13]. La segmentation pourra être réalisée plusieurs fois par un même observateur et par plusieurs observateurs afin d'évaluer la reproductibilité intra- et interobservateur des paramètres extraits.

Prétraitement

Des étapes de prétraitement sont nécessaires avant le calcul des paramètres de radiomique.

Pour le calcul de certains paramètres de texture et leur comparaison entre des lésions de différents patients, une étape d'interpolation est souvent nécessaire permettant d'obtenir des pixels ou des voxels de même taille et isotropes. Par exemple, si une IRM présente des voxels de $1,0 \times 1,0 \times 3,0$ mm que l'on souhaite rendre isotropes, l'interpolation peut être réalisée par un sous-échantillonnage

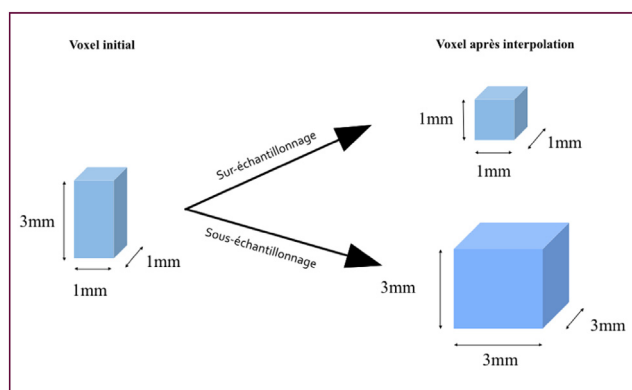


Figure 3. Exemple des deux méthodes possibles d'interpolation d'un voxel de $1,0 \times 1,0 \times 3,0$ mm permettant la création d'un voxel isotrope.

produisant des pixels de $3,0 \times 3,0 \times 3,0$ mm (mais entraînant une perte d'information) ou un sur-échantillonnage produisant des pixels de $1,0 \times 1,0 \times 1,0$ mm (mais nécessitant l'introduction d'informations supplémentaires artificiellement produites) (Fig. 3). Cette étape peut influencer l'analyse de texture et actuellement aucune recommandation ne permet de choisir entre un sous-échantillonnage ou un sur-échantillonnage [11,12].

Avant l'extraction de données quantitatives de texture (second ordre), une autre étape nécessaire est la discrétisation en niveaux de gris. Elle consiste à diminuer et standardiser le nombre de niveaux de gris des images segmentées de la cohorte en regroupant les niveaux de gris proches. On peut ainsi passer de plusieurs milliers de niveaux de gris (en IRM par exemple, où les intensités ne sont pas standardisées) à un nombre plus restreint autorisant le calcul (complexe) de paramètres de texture, leur comparaison d'un patient à l'autre et jouant aussi un rôle de débruitage de l'image.

Deux méthodes principales de discrétisation sont décrites : relative ou absolue.

Discrétisation relative (fixed bin number, FBN)

Celle-ci consiste à fixer un nombre x de niveaux de gris (NG). Toute image sera alors ré-échantillonnée en x niveaux de gris quel que soit le nombre de niveaux de gris de départ.

Par exemple, si une image présente un niveau d'intensité de signal minimum de 10 et un niveau maximum de 330, elle contient 320 niveaux d'intensité de signal. Pour réaliser une discrétisation relative de huit niveaux de gris (FBN8) il convient de répartir les 320 niveaux d'intensité en huit niveaux de $320/8=40$ unités de longueur. Les bornes de niveaux de gris seront donc 10, 50, 90, 130, 170, 210, 250, 290, 330.

Si on applique la même discrétisation à une image ayant une gamme de niveaux de gris différente, il y aura toujours huit niveaux de gris mais leur largeur (pas d'échantillonnage) sera différente.

Discrétisation absolue (fixed bin width, FBW)

Celle-ci consiste à fixer les bornes d'échantillonnage pour toutes les images de la cohorte. La largeur des groupes d'échantillonnage est donc la même pour l'ensemble des examens et le nombre de niveaux de gris, lui, varie suivant le nombre d'intensités de niveaux de gris de départ de l'image.

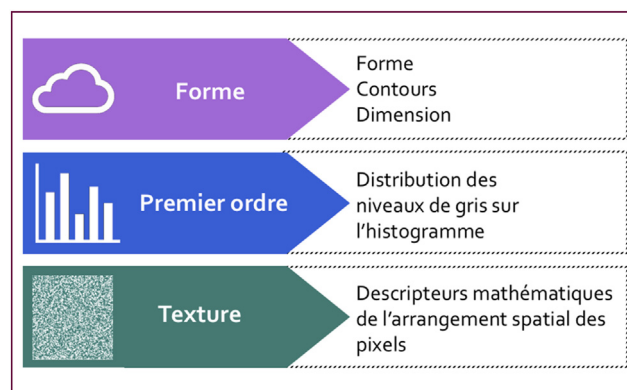


Figure 4. Les trois catégories de paramètres utilisées en radio-mique.

Par exemple, pour la même image présentant un niveau minimum d'intensité de 10 et un niveau maximum d'intensité de 330 une discrétisation absolue de 10 (FBW10) peut être réalisée. Les bornes de niveaux de gris seront dans ce cas de 0, 10, 20, 30, 40, etc. Il y aura donc 32 niveaux de gris.

Si on applique la même discrétisation à une image ayant une gamme d'intensité de niveaux de gris différents, le nombre de niveaux de gris final ne sera pas le même, mais la largeur sera la même.

Le choix de la discrétisation en niveaux de gris a un impact sur les valeurs des paramètres calculés et leur reproductibilité [14–19] ; mais dans la dernière version de l'Initiative de standardisation des biomarqueurs radiologiques (IBSI) datant de septembre 2018, les recommandations laissent libre choix entre les deux types de discrétisations pour les études réalisées en scanner et en TEP-TDM (3). L'IBSI recommande pour l'IRM une discrétisation relative (FBN) qui préserverait mieux le contraste. Néanmoins la seule étude ayant été publiée à ce jour évaluant la discrétisation en IRM conclut que la discrétisation absolue (FBS) permet une meilleure corrélation entre les paramètres de texture calculés et l'évaluation texturale visuelle du radiologue [20].

Extraction des données

Les paramètres extraits sont de trois natures : de forme, de premier ou de second ordre (texture) (Fig. 4).

Paramètres de forme (shape)

Ils décrivent les contours, la forme et les dimensions d'une lésion en deux ou trois dimensions indépendamment des niveaux de gris présents dans l'image. Ils dépendent fortement de la segmentation.

Par exemple : surface de la région d'intérêt (region of interest, ROI) en millimètres carrés, plus grand axe, sphéricité, compacité, allongement, etc.

Paramètres de premier ordre (first order)

Ils décrivent la distribution de l'intensité des niveaux de gris des pixels ou voxels, basée sur l'histogramme.

Par exemple : intensité maximale, intensité moyenne, déviation standard (variance autour de la valeur moyenne), énergie (uniformité de l'histogramme), etc.

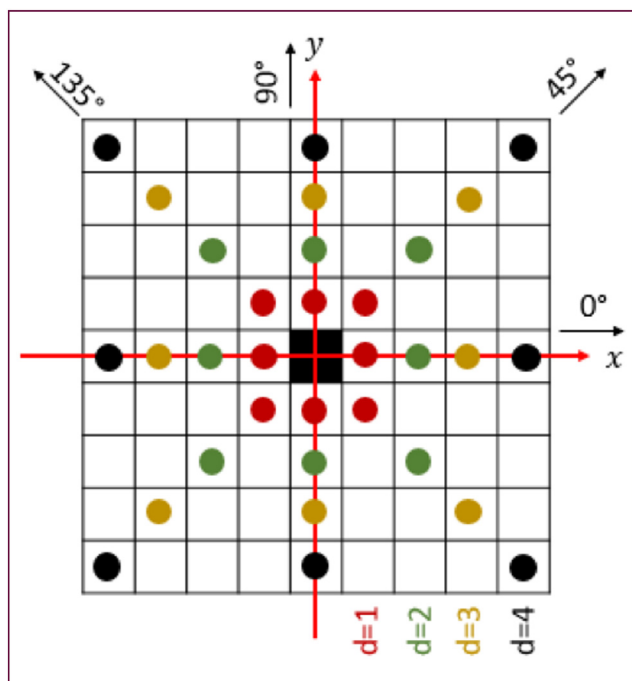


Figure 5. Dans une image bidimensionnelle, un pixel (ici central) peut avoir un maximum de huit voisins en considérant un décalage angulaire θ égal à $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ pour chaque distance d égale à $\{1, 2, 3, 4\}$.

Paramètres de texture (de second ordre)

Ils décrivent la situation des pixels les uns par rapport aux autres (organisation spatiale). Plusieurs méthodes de description existent, produisant chacune un certain nombre de paramètres :

- la matrice de co-occurrence des niveaux de gris (grey-level cooccurrence matrix, GLCM) ;
- la taille des zones homogènes de niveaux de gris (grey-level size zone matrix, GLSZM) ;
- la différence des niveaux de gris de voisinage (neighborhood grey-tone difference matrix, NGTDM) ;
- la longueur de plages de niveau de gris (grey-level run length matrix, GLRLM) ;
- la dépendance en niveaux de gris (grey-level dependence matrix, GLDM).

Par exemple, décrite par Haralick en 1973, la matrice de co-occurrence (GLCM) est une des méthodes d'analyse de texture les plus utilisées [21]. La matrice de co-occurrence est une matrice carrée de taille $N_g \times N_g$, où N_g est le nombre de niveaux de gris dans l'image. Elle représente le nombre d'occurrences de couple de pixels donnés séparés d'une distance d et selon un angle θ (Fig. 5). Chaque valeur de la matrice est calculée en considérant la probabilité de transition d'un pixel d'intensité i à un pixel d'intensité j selon une direction (horizontale, verticale, obliques) définie par l'angle θ et une distance d (Fig. 6). Pour les images bidimensionnelles, chaque pixel peut avoir un maximum de huit voisins : en haut, en bas, à gauche, à droite, en haut à gauche, en haut à droite, en bas à gauche et en bas à droite en considérant une direction " θ " égale à $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ et une distance d égale à $\{1, 2, 3, 4\}$ (Fig. 5). De multiples paramètres sont ensuite calculés à partir de cette matrice comme le cluster shade qui mesure l'asymétrie et l'uniformité de la matrice de co-occurrence, ou le cluster tendency qui est une mesure des groupements de pixels ayant une valeur de niveaux de gris similaires ou encore

la différence d'entropie qui mesure la variabilité des différences d'intensité des niveaux de gris voisins. Cependant, la signification visuelle de chacun de ces paramètres est incertaine.

De nombreux logiciels existent permettant l'extraction des différents paramètres. Qu'ils soient en accès libre (Pyradiomics [22], LifeX [23], etc.) ou « maison » (logiciels souvent dérivés de boîtes à outils Matlab), il apparaît important d'unifier la manière de coder ces paramètres en utilisant les définitions des paramètres de texture telles que décrites dans l'initiative de standardisation des biomarqueurs radiologiques et cela afin de permettre une répétabilité et une uniformisation des études de radiomique [24].

Sélection des paramètres

Les logiciels d'analyse de texture permettent l'extraction de centaines de paramètres par image. Le plus souvent, il est nécessaire de réaliser une réduction du nombre de paramètres afin d'éviter un surapprentissage (overfitting) lié à un nombre insuffisant de patients. Un modèle surajusté est un modèle statistique qui contient plus de paramètres que ne peut le justifier le nombre d'observations, résultant en de « fausses découvertes » liées au hasard statistique. De multiples méthodes de réduction du nombre de paramètres existent et aucune recommandation actuelle ne permet de trancher pour l'une ou pour l'autre mais toutes doivent permettre d'obtenir idéalement dix fois moins de paramètres analysés que d'événements positifs. Par exemple, dans une étude analysant des paramètres quantitatifs extraits d'IRM mammaires présentant une cohorte de 150 patientes, dont 100 atteintes d'un cancer du sein (événement positif), la sélection de paramètres devrait de ne conserver dans le modèle final que dix paramètres ou moins. C'est une condition nécessaire pour limiter le risque de surapprentissage.

Ces méthodes peuvent être classées en trois grandes catégories décrites ci-après.

Méthodes de filtrage

Elles permettent de sélectionner les paramètres en fonction de certaines de leurs caractéristiques, par exemple :

- la robustesse/répétabilité d'une mesure peut être étudiée par la stabilité lors de trois segmentations différentes, les paramètres non reproductibles ou non répétables étant éliminés.
- la non redondance : lorsque plusieurs paramètres sont identifiés comme redondants, il est possible de ne conserver qu'un seul de ces paramètres.

Méthodes de transformation des paramètres

Au lieu de sélectionner une caractéristique par rapport à une autre, ces méthodes créent une nouvelle variable à partir d'une combinaison de plusieurs variables d'intérêt. Par exemple, l'analyse en composante principale crée de nouvelles variables composées de variables initiales afin de décrire au mieux la population.

Méthodes incluses dans le modèle de classification

Les paramètres sont sélectionnés directement selon leur capacité discriminative dans le modèle de classification choisi (c'est le cas des modèles d'apprentissage profond utilisant les réseaux de neurone, par exemple).

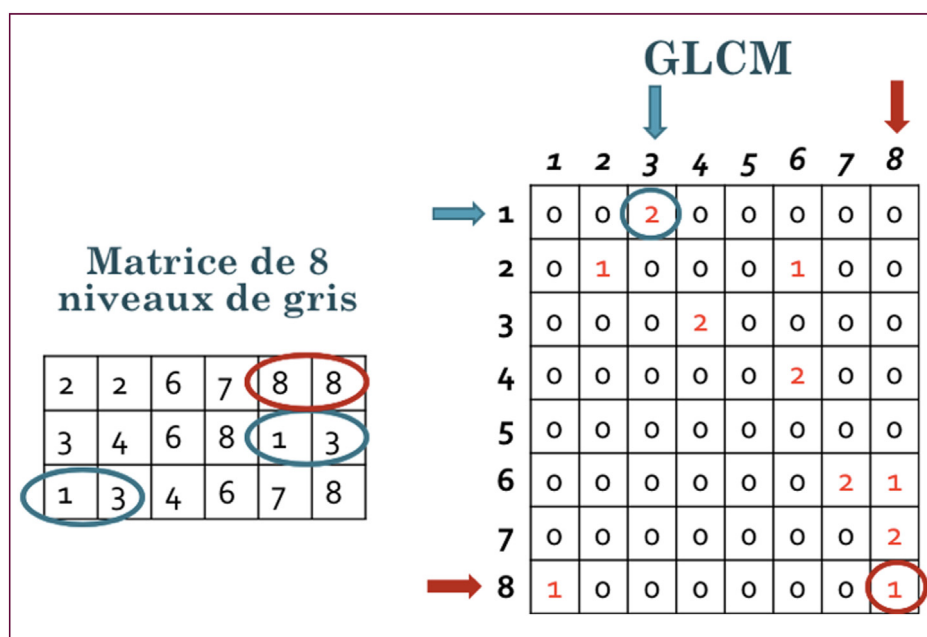


Figure 6. Calcul de la matrice de cooccurrence (GLCM) pour une matrice en huit niveaux de gris avec un angle θ de 0° (horizontal) et une distance d de 1. L'association de pixels de niveaux de gris 1 et 3 est retrouvée deux fois (bleu). L'association de pixels de niveaux de gris 8 et 8 est retrouvée une fois (rouge).

Classification

Les méthodes de classification ont pour but d'entraîner un algorithme permettant d'identifier l'appartenance ou non d'un individu à un groupe (présence ou absence de maladie, lésion bénigne ou maligne, etc.) en réalisant des prédictions à partir d'un grand nombre de données.

Il existe deux types d'apprentissage automatique, selon que la classification est supervisée ou non.

Classification supervisée

À partir d'individus étiquetés (par exemple bénin/malin) et d'une base de données descriptive contenant les paramètres de radiomique mesurés, l'algorithme est optimisé pour qu'il apprenne les caractéristiques de chaque groupe de patient. Une fois entraîné, l'algorithme sera capable, lorsqu'on lui présente un nouvel individu, de prédire la catégorie du patient. Des exemples de méthodes de classification supervisée sont :

- les séparateur à vastes marges (support vector machine, SVM) permettant de rechercher le plan séparant au mieux, avec la marge la plus importante, les différents groupes de patients ;
- l'arbre de décision créé par répartition récursive qui permet de sélectionner les paramètres les plus pertinents ainsi que leurs valeurs seuil pour classer deux populations. Cette méthode utilisée dans deux articles de Thomassin-Naggara et al. a par exemple permis la classification des lésions bénignes et malignes avec une aire sous la courbe de spécificité/sensibilité (area under the receiver operating characteristic curve, AUC ROC, ou simplement AUC) de 0,90 pour l'étude portant sur les rehaussements non-masses en IRM mammaire et une AUC supérieure à 0,94 pour l'étude portant sur les masses ovariennes [25,26].

Classification non supervisée

Il s'agit de trouver un modèle mathématique regroupant les données les plus proches et séparant les données les plus différentes (on parle de clustering). L'apprentissage est dit non supervisé car les données ne sont pas étiquetées, autrement dit l'algorithme n'a pas de cible pour corréler les données d'entrée et trouver un modèle pertinent. Il doit donc se contenter de trouver des comportements semblables dans les données et apprendre le modèle qui permet de les décrire au mieux. Une fois entraîné, l'algorithme sera capable de classer une nouvelle donnée dans l'un ou l'autre des groupes formés avec une bonne pertinence.

Par exemple, le regroupement par K-moyenne (K-means clustering) crée K regroupements en comparant le degré de similarité des observations. Deux données qui se ressemblent auront une distance de dissimilarité réduite. Dans un tel algorithme, une observation ne peut être présente que dans un seul regroupement.

Dans un article publié en 2015, l'équipe de Parmar a étudié 14 méthodes de sélection de paramètres et 12 méthodes de classification sur 440 paramètres de radiomique extraits de 464 scanners présentant des cancers du poumon [27]. Cette étude a permis d'identifier le couple sélection de paramètre/classification par un test de Wilcoxon et un arbre probabiliste comme la meilleure combinaison dans le cadre de son étude. Une analyse multivariée de variance (Anova) sur les scores AUC était également été réalisée, permettant de montrer que la méthode de classification était la plus grande source de variabilité de l'AUC (34,21 % de la variance des scores AUC contre 6,25 % pour la sélection de paramètres et 23,03 % pour l'interaction entre le classifieur et la méthode de sélection de paramètres).

Validation

Plus le nombre de variables utilisées dans un modèle augmente, plus il décrit précisément la population étudiée mais moins le modèle est généralisable à une autre population.

C'est ce qui est appelé le surapprentissage ou overfitting. Afin de s'assurer que le modèle trouvé est généralisable à une autre population comparable (même modalité, mêmes lésions) il est nécessaire de valider le modèle.

La méthode standardisée de validation est celle réalisée sur un groupe test. La population est séparée dès le début de l'expérience en deux groupes présentant des caractéristiques comparables notamment en termes de prévalence de la maladie étudiée ; un groupe « entraînement », sur lequel sera réalisé et entraîné le modèle de classification et un groupe « test » sur lequel sera appliqué le modèle afin d'en tester les performances.

Lorsque l'on ne dispose pas d'une population suffisante pour réaliser un groupe test, une méthode de validation croisée (cross validation) peut être réalisée afin d'évaluer les performances de l'algorithme sans perdre en quantité de données d'entraînement. Le jeu de données est découpé en k parties qui vont tour à tour être utilisées comme jeu de test. Lorsqu'une partie est utilisée comme jeu de test, le reste des données est utilisé comme entraînement. À la fin de l'expérience, chaque observation a servi une fois dans le jeu test et $k-1$ fois dans le jeu d'entraînement. On obtient une prédiction par partie k et aucune de ces prédictions n'a été réalisée sur un jeu d'entraînement contenant la population test. La performance globale est obtenue en moyennant les résultats de chaque partie.

Avant de pouvoir prétendre à une application clinique, un modèle de classification utilisant des paramètres quantitatifs extraits de l'image doit avoir démontré ses performances diagnostiques sur un ou plusieurs groupes tests issus de centres différents et si possible de machine différentes.

Afin de s'assurer de la qualité d'une étude de radiomique, Lambin et al. ont publié un score de qualité des études de radiomique (radiomics quality score, RQS) notant et détaillant les bonnes pratiques du protocole d'acquisition jusqu'à la mise à disposition des données (images, segmentation, code) permettant une validation externe indépendante [28].

Exemples d'applications

Cette partie n'a pas pour but de réaliser une étude exhaustive de l'ensemble de la littérature de radiomique en imagerie de la femme, mais de présenter certaines études prometteuses ainsi que leurs limites.

IRM mammaire

L'utilisation de paramètres quantitatifs extraits de l'IRM mammaire a été étudiée dès 2004 dans un article de Chang et al. évaluant la réponse thérapeutique à l'aide de paramètres de texture de premier ordre issus de l'histogramme [29]. Depuis de nombreuses équipes se sont intéressées à l'utilisation de la radiomique en IRM mammaire et plusieurs études sont publiées chaque mois à ce sujet.

Certaines études se sont intéressées à la prédiction du diagnostic histologique obtenu après biopsie ou chirurgie. Dans une étude publiée en 2010 par Bhooshan et al., l'analyse par un réseau bayésien de 353 paramètres issus d'IRM mammaire permettait de différencier un carcinome canalaire in situ d'un carcinome canalaire invasif avec une AUC de $0,83 \pm 0,03$, un carcinome canalaire invasif d'une lésion bénigne avec une AUC de $0,82 \pm 0,002$ et un carcinome canalaire in situ d'une lésion bénigne avec une AUC de

$0,79 \pm 0,03$ [30]. D'autres études comme celles publiées par l'équipe de Fan et al. se sont intéressées à la prédiction de marqueurs biologiques tumoraux. Dans une première étude publiée en 2017 portant sur 60 cancers du sein, un modèle permettait de prédire avec une AUC de 0,888 l'expression des récepteurs de l'épidermal growth factor (Human epidermal growth factor receptor-2, HER2) par la lésion [31]. Dans une deuxième étude de la même équipe publiée en 2018 portant sur 77 cancers du sein dont la texture était analysée en sous-régions, un modèle permettait la prédiction de l'expression du ki-67 avec une AUC de 0,807 [32].

De nombreuses études se sont intéressées à la capacité de prédiction de la réponse thérapeutique par la radiomique [29,30,33–39]. Dans une étude rétrospective récente publiée par l'équipe de Brahman et al. portant sur 117 patientes bénéficiant d'une chimiothérapie néoadjuvante, l'extraction et l'analyse de 99 paramètres de texture sur l'IRM préthérapeutique permettait la prédiction de la réponse complète (T0 ou Tis chirurgical) avec une AUC de $0,78 \pm 0,030$ [40]. Lorsque l'expression des récepteurs hormonaux et de l'EGF étaient ajoutés aux données de texture, la prédiction de la réponse à la chimiothérapie néoadjuvante atteignait une AUC de $0,93 \pm 18$.

Enfin, certaines études se sont aussi penchées sur l'évaluation du pronostic des patientes en faisant le lien entre génomique et radiomique. C'est le cas de l'étude de l'équipe de Li et al. s'intéressant à la corrélation entre des résultats d'une analyse radiomique issue de 84 IRM mammaires réalisées chez des patientes atteintes de cancer du sein invasif et le résultat de trois tests génétiques pronostiques qui étaient l'Oncotype DX, le MammaPrint et le PAM50 [41]. L'analyse IRM permettait de prédire le résultat des tests génétiques avec une AUC de 0,88 pour le MammaPrint, 0,76 pour l'Oncotype DX, 0,68 pour le PAM50 basé sur le sous-type et 0,55 sur le PAM50 basé sur le sous-type et la prolifération.

Cancers gynécologiques

Plus récemment, plusieurs équipes ont étudié l'intérêt de la radiomique dans l'analyse notamment pronostique des cancers gynécologiques.

Comme l'ont décrit récemment Nougaret et al., des marqueurs issus de la radiomique décrivant l'hétérogénéité tumorale en imagerie apparaissent comme intéressants pour le pronostic des cancers ovariens, au même titre que l'hétérogénéité moléculaire étudiée en génomique [42–47]. Dans une étude récente de Thomassin-Naggara et al. portant sur 113 masses ovariennes (52 bénignes et 61 malignes) analysées en IRM, l'augmentation de la déviation standard de l'intensité du signal (paramètres de premier ordre caractérisant l'hétérogénéité tumorale) dans les 90 premières secondes après injection de chélates de gadolinium permettait de prédire la malignité tumorale avec une sensibilité de 75,4 % et une spécificité de 80,8 % [48]. Rizzo et al. ont également montré en étudiant 516 paramètres extraits de tumeurs ovariennes segmentées manuellement en trois dimensions sur les scanners préthérapeutiques de 101 patientes que l'ajout des paramètres F2 shape et Max3DDiameter permettait d'améliorer la prédiction clinique de la progression tumorale à 12 mois de 14 % (AUC de 0,87 contre 0,73) [49]. Cette étude prometteuse présentait comme principale limite son absence de validation, le modèle ayant été évalué sur le même groupe de patients que celui ayant permis l'entraînement, sans groupe test ni validation croisée.

Dans une étude de Ueno et al. portant sur 137 patientes atteintes d'un carcinome de l'endomètre, l'utilisation dans un modèle d'arbre aléatoire (random forest) de paramètres extraits de plusieurs séquences IRM (T2, diffusion et après injection de chélates de gadolinium) segmentées manuellement permettait de diagnostiquer l'envahissement myométrial profond avec une AUC de 0,84 (intervalle de confiance [IC] : 0,69–0,90), l'envahissement lymphatique avec une AUC de 0,80 (IC : 0,74–0,87) et les types histologiques de haut-grade avec une AUC de 0,83 (IC : 0,76–0,89) [50].

Une étude multicentrique canadienne et française, publiée par l'équipe de Lucia et al. [51] en décembre 2018 utilisant 112 patients dans le groupe d'entraînement et 78 dans le groupe test et portant sur la prédiction de l'évolution après la chimiothérapie des cancers du col, montrait que l'utilisation de deux paramètres de texture (GLNU_{GLRLM} issu de la TEP-TDM et Entropy_{GLCM} issu d'une carte de coefficient apparent de diffusion) permettait de prédire la survie sans récurrence avec une précision de 90 % (IC : 79–98 %), une sensibilité de 92 à 93 % et une spécificité de 87 à 89 % ainsi que le contrôle locorégional avec une précision de 98 % (90–99 %), une sensibilité de 86 % et une spécificité de 100 %. Cette étude permettait également de discuter de la généralisation des études de radiomique à des cohortes multicentriques. En effet, plusieurs études ont montré la sensibilité des paramètres de radiomique aux changements de machine ou de paramètres de reconstruction [52–54]. Les résultats de cette étude suggèrent qu'une harmonisation statistique des résultats de textures obtenus dans des centres différents grâce à une méthode nommée ComBat empruntée à la génomique permettrait une amélioration de la généralisation précision des résultats [55].

Conclusion

La radiomique est une méthode prometteuse permettant, à partir d'un très grand nombre de paramètres quantitatifs extraits d'images radiologiques et confrontés à la réalité clinique et biologique, de découvrir de nouveaux biomarqueurs diagnostiques, pronostiques ou prédictifs. Son application à l'imagerie de la femme et notamment en IRM mammaire pourrait permettre une avancée dans la prise en charge diagnostique et thérapeutique des patientes. Néanmoins, malgré des centaines de publications chaque année, l'application clinique de ces études reste encore très limitée. La généralisation de leurs résultats demande l'utilisation d'une méthodologie stricte et détaillée, d'une validation externe et la création d'outils dédiés accessibles depuis une console de post-traitement utilisable en clinique.

Déclaration de liens d'intérêts

Les auteurs déclarent ne pas avoir de liens d'intérêts.

Références

- [1] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGP, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48(4):441–6.
- [2] van Helden P. Data-driven hypotheses. *EMBO Rep* 2013;14(2):104.
- [3] Zwanenburg A, Leger S, Vallières M, Löck S. Initiative for the IBS. Image biomarker standardisation initiative. *ArXiv161207003 Cs* [Internet] 2016 <http://arxiv.org/abs/1612.07003>.
- [4] Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30(9):1234–48.
- [5] Cai L, Wang X, Wang Y, Guo Y, Yu J, Wang Y. Robust phase-based texture descriptor for classification of breast ultrasound images. *Biomed Eng OnLine* [Internet] 2015;14:26.
- [6] Chen C-H, Chang C-K, Tu C-Y, Liao W-C, Wu B-R, Chou K-T, et al. Radiomic features analysis in computed tomography images of lung nodule classification. *PLoS ONE* 2018;3(2):e0192002.
- [7] Mao N, Yin P, Wang Q, Liu M, Dong J, Zhang X, et al. Added value of radiomics on mammography for breast cancer diagnosis: a feasibility study. *J Am Coll Radiol JACR* 2018;041, <http://dx.doi.org/10.1016/j.jacr.2018.09.041>.
- [8] Cai H, Peng Y, Ou C, Chen M, Li L. Diagnosis of breast masses from dynamic contrast-enhanced and diffusion-weighted MR: a machine learning approach. *PLoS One* 2014; 2014;9(1):e87387.
- [9] Gudbjartsson H, Patz S. The rician distribution of noisy MRI data. *Magn Reson Med* 1995;34(6):910–4.
- [10] Gjesteby L, Man BD, Jin Y, Paganetti H, Verburg J, Giantsoudi D, et al. Metal artifact reduction in CT: where are we after four decades? *IEEE* 2016;4:5826–49.
- [11] Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage* 2006;31(3):1116–28.
- [12] Owens CA, Peterson CB, Tang C, Koay EJ, Yu W, Mackin DS, et al. Lung tumor segmentation methods: Impact on the uncertainty of radiomics features for non-small cell lung cancer. *PLoS One* 2018;13(10):e0205003.
- [13] Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust Radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One* 2014;9(7):e102107.
- [14] Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 2017;44(3):1050–62.
- [15] Altazi BA, Zhang GG, Fernandez DC, Montejo ME, Hunt D, Werner J, et al. Reproducibility of (¹⁸F)-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms. *J Appl Clin Med Phys* 2017;18(6):32–48.
- [16] Desseroit M-C, Tixier F, Weber WA, Siegel BA, Chèze Le Rest C, Visvikis D, et al. Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non-small cell lung cancer tumors: a repeatability analysis in a prospective multicenter cohort. *J Nucl Med* 2017;58(3):406–11.
- [17] Hatt M, Majdoub M, Vallières M, Tixier F, Le Rest CC, Groheux D, et al. (¹⁸F)-FDG.P.E.T uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med* 2015;56(1):38–44.
- [18] Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging* 2017;44(1):151–65.
- [19] Leijenaar RTH, Nalbantov G, Carvalho S, van Elmpt WJC, Troost EGC, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep* 2015;5:11075.
- [20] Goya-Outi J, Orhac F, Calmon R, Alentorn A, Nioche C, Philippe C, et al. Computation of reliable textural indices from multimodal brain MRI: suggestions based on a study of patients with diffuse intrinsic pontine glioma. *Phys Med Biol* 2018;63(10):105003.

- [21] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;SMC-3(6):610–21.
- [22] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77(21):e104–7.
- [23] Nioche C, Orhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res* 2018;78(16):4786–9.
- [24] Zwanenburg A, Leger S, Vallières M, Löck S. Initiative for the IBS. Image biomarker standardisation initiative. *ArXiv161207003 Cs* 2016 <http://arxiv.org/abs/1612.07003>.
- [25] Thomassin-Naggara I, Aubert E, Rockall A, Jalaguier-Coudray A, Rouzier R, Daraï É, et al. Adnexal masses: development and preliminary validation of an MR imaging scoring system. *Radiology* 2013;267(2):432–43.
- [26] Thomassin-Naggara I, Trop I, Chopier J, David J, Lalonde L, Daraï É, et al. Nonmasslike enhancement at breast MR imaging: the added value of mammography and US for lesion categorization. *Radiology* 2011;261(1):69–79.
- [27] Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 2015;5:13087.
- [28] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14(12):749–62.
- [29] Chang Y-C, Huang C-S, Liu Y-J, Chen J-H, Lu Y-S, Tseng W-YI. Angiogenic response of locally advanced breast cancer to neoadjuvant chemotherapy evaluated with parametric histogram from dynamic contrast-enhanced MRI. *Phys Med Biol* 2004;49(16):3593–602.
- [30] Bhooshan N, Giger ML, Jansen SA, Li H, Lan L, Newstead GM. Cancerous breast lesions on dynamic contrast-enhanced MR images: computerized characterization for image-based prognostic markers. *Radiology* 2010;254(3):680–90.
- [31] Fan M, Li H, Wang S, Zheng B, Zhang J, Li L. Radiomic analysis reveals DCE-MRI features for prediction of molecular subtypes of breast cancer. *PloS One* 2017;12(2):e0171683.
- [32] Fan M, Cheng H, Zhang P, Gao X, Zhang J, Shao G, et al. DCE-MRI texture analysis with tumor subregion partitioning for predicting Ki-67 status of estrogen receptor-positive breast cancers. *J Magn Reson Imaging JMRI* 2018;48(1):237–47.
- [33] Ahmed A, Gibbs P, Pickles M, Turnbull L. Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *J Magn Reson Imaging JMRI* 2013;38(1):89–101.
- [34] Ashraf A, Gaonkar B, Mies C, DeMichele A, Rosen M, Davatzikos C, et al. Breast DCE-MRI kinetic heterogeneity tumor markers: preliminary associations with neoadjuvant chemotherapy response. *Transl Oncol* 2015;8(3):154–62.
- [35] Kim J-H, Ko ES, Lim Y, Lee KS, Han B-K, Ko EY, et al. Breast Cancer heterogeneity: MR imaging texture analysis and survival outcomes. *Radiology* 2017;282(3):665–75.
- [36] Parikh J, Selmi M, Charles-Edwards G, Glendenning J, Ganeshan B, Verma H, et al. Changes in primary breast cancer heterogeneity may augment midtreatment MR imaging assessment of response to neoadjuvant chemotherapy. *Radiology* 2014;272(1):100–12.
- [37] Thibault G, Tudorica A, Afzal A, Chui SY-C, Naik A, Troxell ML, et al. DCE-MRI texture features for early prediction of breast cancer therapy response. *Tomogr J Imaging Res* 2017;3(1):23–32.
- [38] Wu J, Cao G, Sun X, Lee J, Rubin DL, Napel S, et al. Intratumoral spatial heterogeneity at perfusion MR imaging predicts recurrence-free survival in locally advanced breast cancer treated with neoadjuvant chemotherapy. *Radiology* 2018;288(1):26–35.
- [39] Wu J, Gong G, Cui Y, Li R. Intratumor partitioning and texture analysis of dynamic contrast-enhanced (DCE)-MRI identifies relevant tumor subregions to predict pathological response of breast cancer to neoadjuvant chemotherapy. *J Magn Reson Imaging JMRI* 2016;44(5):1107–15.
- [40] Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res BCR* 2017;19(1):57.
- [41] Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. *Radiology* 2016;281(2):382–91.
- [42] Nougaret S, Tardieu M, Vargas HA, Reinhold C, Vande Perre S, Bonanno N, et al. Ovarian cancer: An update on imaging in the era of radiomics. *Diagn Interv Imaging* 2018, <http://dx.doi.org/10.1016/j.diii.11.007>.
- [43] Bai H, Cao D, Yang J, Li M, Zhang Z, Shen K. Genetic and epigenetic heterogeneity of epithelial ovarian cancer and the clinical implications for molecular targeted therapy. *J Cell Mol Med* 2016;20(4):581–93.
- [44] Cancer Genome Atlas ResearchNetwork. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;474(7353):609–15.
- [45] Konecny GE, Winterhoff B, Wang C. Gene-expression signatures in ovarian cancer: Promise and challenges for patient stratification. *Gynecol Oncol* 2016;141(2):379–85.
- [46] Lambrechts S, Smeets D, Moisse M, Braicu EI, Vanderstichele A, Zhao H, et al. Genetic heterogeneity after first-line chemotherapy in high-grade serous ovarian cancer. *Eur J Cancer* 2016;53:51–64.
- [47] Patch A-M, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S, et al. Whole-genome characterization of chemoresistant ovarian cancer. *Nature* 2015;521(7553):489–94.
- [48] Thomassin-Naggara I, Soualhi N, Balvay D, Daraï É, Cuénod C-A. Quantifying tumor vascular heterogeneity with DCE-MRI in complex adnexal masses: A preliminary study. *J Magn Reson Imaging JMRI* 2017;46(6):1776–85.
- [49] Rizzo S, Botta F, Raimondi S, Origgi D, Buscarino V, Colarieti A, et al. Radiomics of high-grade serous ovarian cancer: association between quantitative CT features, residual tumour and disease progression within 12 months. *Eur Radiol* 2018;28(11):4849–59.
- [50] Ueno Y, Forghani B, Forghani R, Dohan A, Zeng XZ, Chamming's F, et al. Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification-a preliminary analysis. *Radiology* 2017;284(3):748–57.
- [51] Lucia F, Visvikis D, Vallières M, Desseroit M-C, Miranda O, Robin P, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. *Eur J Nucl Med Mol Imaging* 2018, <http://dx.doi.org/10.1007/s00259-018-4231-9>.
- [52] Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, et al. Impact of Image reconstruction settings on texture features in (18F)-FDG PET. *J Nucl Med* 2015;56(11):1667–73.
- [53] Reuzé S, Orhac F, Chargari C, Nioche C, Limkin E, Riet F, et al. Prediction of cervical cancer recurrence using textural features extracted from (18F)-FDG PET images acquired with different scanners. *Oncotarget* 2017;8(26):43169–79.
- [54] Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, et al. Harmonization of multi-site diffusion tensor imaging data. *NeuroImage* 2017;161:149–70.
- [55] Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat* 2007;8(1):118–27.