



Current Advances in Computational biology of Lung Cancer

Purbasha Pati^{1,2}, Pooja Kumari^{1,3}, Neha Kumari^{1,4}, Kumari Surekha Mahto^{1,5}, Shekhar Marandi^{1,5}, Sadaf Naaz^{1,2}, Sweety Guria Rani^{1,2}, Naiyar Naaz², Rupa Verma⁵, Saurabh Kumar Gupta¹, Mukesh Nitin¹

¹Department of Tech. Biosciences, Digianalix, South Samaj Street, Tharpakhna, Ranchi, Jharkhand, INDIA.

²Department of Biotechnology, Ranchi Women's College, Circular Rd. Nagra Toli, Ranchi, Jharkhand, INDIA.

³Department of Applied Microbiology, Vellore Institute of Technology, Vellore, Tamil Nadu, INDIA.

⁴Department of Biotechnology, Birla Institute of Technology, Mesra, Jharkhand, INDIA.

⁵Department of Biotechnology, Ranchi University, Ranchi, Jharkhand, INDIA.

ABSTRACT: Lung cancer is the main basis of cancer death amongst men and women, making up almost 25% of the world's total cancer deaths. Lung cancer described for nearly 1.6 million deaths in 2012 and 1.80 million deaths in 2020. Small cell lung cancer and non-small-cell lung cancer are the two key categories of Lung cancer. The signs of lung cancer include hemoptysis, weight loss, shortness of breath and chest pain. Lung cancer treated by chemotherapy, surgery and CT scan. In this review paper, one of the most crucial zones aiming lung cancer diagnosis has been discussed. Computer-aided diagnosis (CAD) systems adapted for lung cancer can increase the patients' survival chances. A typical CAD system for lung cancer functions in the fields of lung segmentation, detecting lung nodules and the diagnosis of the nodules as benign or malignant. CAD systems for lung cancer are examined in a huge number of research case studies. CAD system steps and outlining of inhibitor genes at molecular level is being discussed. An insight into multi-omics and molecular dynamics simulations is also given in this paper.

KEYWORDS: Benign or malignant nodule diagnosis, CAD system for lung cancer, Lung cancer diagnosis, Multi-omics, MD simulations.

1. INTRODUCTION

Lung cancer can benign and malignant lung tumor caused due to uncontrolled cell growth in the lung tissues. [1,2] The major countries like North America, Europe, and East Asia, along with over one-third of new cases in China are reported of highest number of patients subjected to lung cancer. While the African and South Asian countries are reported much lower [3]. Globally, in 2012, lung cancer occurred in 1.8 million people and resulted in 1.6 million deaths [4]. The Malaysian National Cancer Registry (2007) reported 1865 lung cancer cases with males 70.8% and females 29.2% [5]. Contributory factors also include long-term tobacco smoking (85%); 10 to 15% of cases have also occurred in people who have never smoked [6]. The other factors include exposure to random gas, asbestos, second-hand smoke, or other forms of air pollution [7]. The two main subcategories of lung cancer are small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma (NSCLC) [8]. NSCLC stage with 80-85% are broadly categorized into three sub-categories, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. These three categories of lung carcinoma are broadly subjected to research studies. However, some other sub-forms of lung carcinoma are also there like adenosquamous and sarcomatoid carcinoma. The SCLC are about 10-15% of all lung cancer and also known as OAT CELL CANCER. The most common clinical symptoms are haemoptysis, weight loss, shortness of breath and chest pain [9]. This disease is detected by chest radiographs, computed tomography scans and Computer-aided diagnosis as well. The diagnosis is confirmed with biopsy by bronchoscopy or CT guidance [10]. Standard treatment includes surgery, chemotherapy, and radiography [11]. NSCLC may from time to time treated through surgery, although chemotherapy and radiotherapy aid in the dealing of SCLC [12]. In 2012, 1.82 million nascent cases were reported and 1.56 million deaths due to lung cancer worldwide [13]. People over 50 years of age having a history of smoking are likely to develop lung cancer. The mortality rate in men has declined in terms of women's lung cancer mortality rate [14]. In the various research studies conducted in USA, the foremost lifetime risk of lung cancer development is 0.8% in men and 0.6% in women are reported so far [15]. The genetic damage to DNA and epigenetic changes results in cell proliferation, apoptosis, and DNA repair of cells, increasing the risk of cancer [16]. Pathogenesis of lung cancer involves activating oncogenes or inactivation of tumor suppressor genes in the lungs [17]. Because mutations occurs within these genes, which in turn induce the development of cancer [18]. So mutations in the K-ras proto-

oncogene is responsible for inducing 10-30% of lung adenocarcinomas [19]. About 0.4% of non-small-cell lung carcinomas involve an EML4-ALK tyrosine kinase fusion gene [20]. DNA methylation, histone tail modification, microRNA regulation led to the inactivation of tumor suppressor genes. [21]. Epidermal growth factor receptor (EGFR) regulates the cell proliferation, apoptosis, angiogenesis, and tumor invasion [22]. Mutations and amplification of EGFR are most frequently seen in non-small-cell lung carcinoma and such modifications are treated with EGFR-inhibitors [23]. Other genes that are also reported to be commonly mutated or amplified are KRAS and BRAF [24].

2. COMPUTER-AIDED DIAGNOSIS OF LUNG CANCER

The developmental research of Computer aided Diagnosis (CAD) systems, several specialized approaches were done in a systematized manner. The primary pace involves the segmentation of lung tissues on chest images to reduce the search space for lung nodules. They were followed by the detection and segmentation of lung nodules from the available search space. The detected nodules were assigned into benign and malignant in the final step. Classification of the detected nodules is a significant component in CAD systems for detecting and diagnosing lung nodules in CT scans. CAD comprises a classification component for detection that categorizes the nodule applicants to nodules or non- nodules, whereas a CAD system for diagnosis classifies detected nodules into benign or malignant nodules. Further, thorough process in developing CAD systems, lung segmentation, nodule detection, nodule segmentation, and nodule diagnosis are addressed.

2.1. Lung Segmentation:

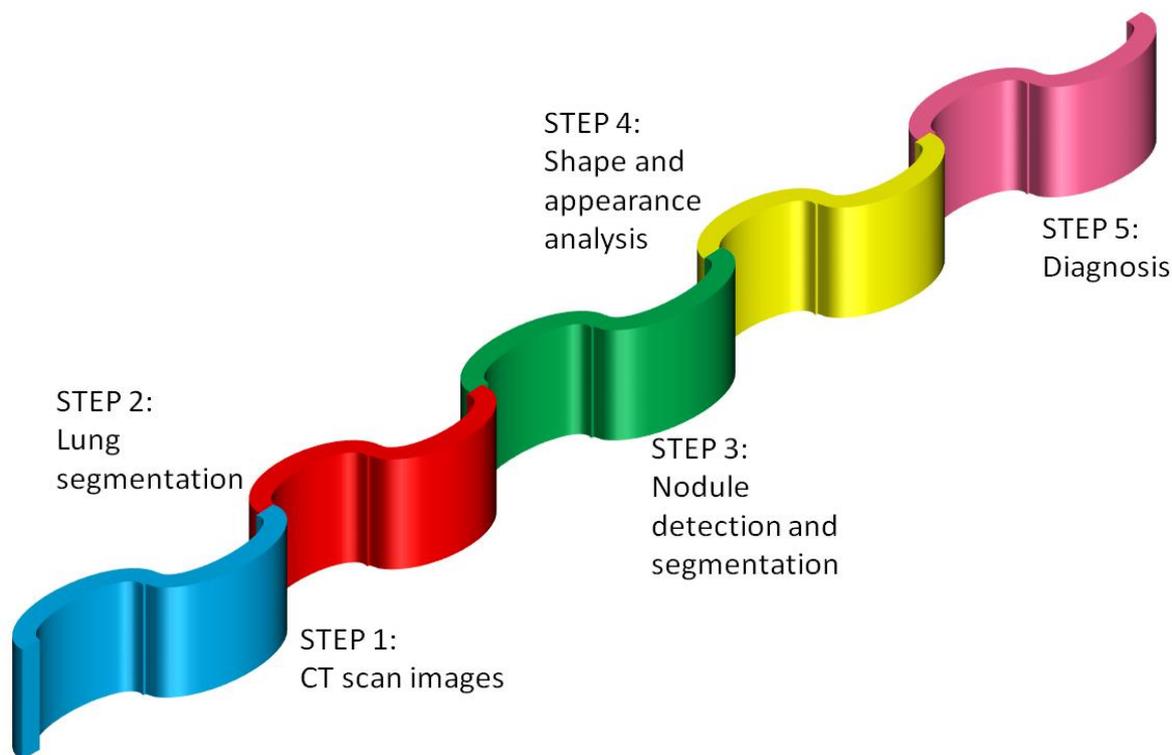


Figure 1. The categorical representation of lung and nodule segmentation approach under computer aided diagnosis for lung cancer

The lung and nodule segmentation steps demonstrated in Figure 1, shows a typical CAD system for lung cancer. The mentioned figure above consists of 5 steps which will result into the integration of CAD system.

Preliminary diagnosis of lung cancer was carried out by the segmentation of lung tissues on the CT scan images. It was considered a vital step in the CAD system. The threshold was computed to get an initial lung region [25]. Then the initial segmentation is



refined by opening and closing morphological operations. Similar method was also adapted for lung lobe segmentation [26, 27]. A threshold-based segmentation approach based on removal of a large airway from the lung region using isotropic diffusion and morphological smoothing of the lung boundary [28]. A graph-based search incorporating intensity is used for segmenting lung fields using a gradient, boundary smoothness and rib information [29]. However, a repetitive approach of Markov-Gibbs-random-field (MGRF) segmentation was used to segment the lung fields from LDCT images [30-32].

One of the biggest challenges for CAD systems is to discriminate the true nodules from vessels, ribs. CAD was monitored in terms of diagnostic accuracy, speed, and automation level. CAD systems for detecting lung nodules in thoracic CT scan covered in two major steps: (i) primary samples were standardly preserved. The segmentation of nodule samples within the lungs was carried out using a fuzzy clustering approach [33,34]. The discrimination between nodule regions and typical structures was observed using the maximum distance inside connected components (MDCC) in 3D X-ray CT images [35]. The standard surface overlap (SNO) method was used in the detection of lung nodules and colon polyps [36]. This method describes the shape and geometry of a potential nodule. Above mentioned overview demonstrates the importance of investigation in designing CADe systems to diagnose lung nodules, including speed, automation level, and ability to see different sizes of nodules. The nodule segmentation is an essential and crucial step in lung cancer applications.

2.2. Lung Nodule Segmentation:

Lung nodule segmentation delineates the nodular lesions in chest CT scans, providing a critical foundation of CAD of lung cancers [37-39]. Under this section, the study on the clinical role of lung nodule segmentation was carried out.

2.2.1. Clinical applications:

Accurate nodule segmentation is crucial for various diagnostic and treatment procedures of lung cancer [40], such as monitoring tumor response against therapy [41, 42], computer-aided lung cancer screening for early detection [43-45], computer-aided diagnosis of tumor malignancy [46, 47] is carried out at high accuracy during nodule segmentation for better diagnosis. Lung cancer screening by CADe enhances the detection of segmented nodules [48].

2.2.2. CT Segmentation techniques:

Technical approaches for volumetric lung nodule segmentation suggests that it is classified into 11 categories which are listed below:

- (i) **Thresholding (TH):** It yields a binary segmentation of the volume of interest by labelling each voxel by testing whether its intensity value surpasses a specific threshold value or not [49].
- (ii) **Mathematical Morphology (MM):** It is a technique in lung nodule segmentation used for handling cases attached to non-target structures such as vessels and the parenchymal wall or the diaphragm. It makes available four basic operators: - opening, closing, erosion, dilation.
- (iii) **Region Growing (RG):** It discovers the connected-component region covering the total seeded pixel by adding next nearby pixels, which is responsible for logical pixel intensity [50].
- (iv) **Deformable Model (DM):** The model represents segmentation methods based on the iterative evolution of contour curves modelling the boundary of a target object, for example typical energy-minimization-based active contour (AC) [51], edge-based geodesic AC [52] and region-based variational LS [53].
- (v) **Dynamic Programming (DP):** DP states to energy minimization variations for detecting optimal contours in images [54].
- (vi) **Spherical/Ellipsoidal Model Fitting:** This fitting exploits the proximity of CT scan lung nodule appearance to the standard Gaussian intensity model. Using approximation models for CADe of nodules in early days.
- (vii) **Probabilistic classification (PC):** This explains each voxel is termed as nodules or other structure. At each voxel, the classification of standard estimation framework, such as MAP, maximum likelihood (ML), and likelihood ratio test is cast [56].
- (viii) **Discriminative classification (DC):** In this it casts the segmentation problem as a voxel-wise classification similar to PC. There exist numerous methods for supervised discriminative classifiers in the machine learning literature.
- (ix) **Mean shift:** MS is a segmentation approach based on an iterative feature space analysis. The MS algorithm performs by clustering feature data points from each data point and the non-parametric distribution modes were calculated [57].
- (x) **Graph-Cuts (GCs) and Watersheds (WSs):** GCs [58] and WSs [59] are the other known well-techniques of traditional image segmentation these are implemented for nodule segmentation problems. Many Researchers applied GC to derive their initial 2D



nodule segmentation. Reported methods of volumetric segmentation are given in (Table 1).

Table 1. Studies on volumetric segmentation reported from 1998 to 2005.

STUDY TYPE	METHOD	PURPOSE	REFERENCES
3D	Geometric deformation flow of 3D L.S surface.	Solitary, solid	[47]
Threshold [2D]/[3D]	Efficient average gradient computation	Small, Juxta vascular	[66]
Dynamic programming,2D	2D contour optimization by DP.	Juxta vascular	[179]
Mathematical morphology	Isotropic resampling for partial-volume effect	Small, solid	[117]
Gaussian fitting and mean shift (MS)	Robust anisotropic Gaussian intensity and model fitting with MS.	Small, Juxta vascular, GO	[75]
Surface analysis	Lung surface removal of juxta-pleural nodules is done by local surface smoothing.	Juxta pleural	[74]

2.2.3. PET Segmentation techniques:

Positron emission tomography (PET) it has been broadly used in oncology applications such as lung cancer detection and nodule segmentation. The generally used quantifier in PET imaging is the standardized uptake value (SUV) that estimates the intensity of the lesion on PET.

$$SUV = c(t) / [\text{Injected dose}(t)/\text{body weight}]$$

To combine both PET and CT curated data, a fusion technique integrates both PET and CT scan images. The fusion technique is classified into three categories: (i)visual fusion in both imaging modalities, (ii)software fusion, and (iii)hardware fusion.

Steenbakkars [60] and Fox [61] used a software fusion method and examined the observer variations in 2 phases, one with CT scans and another with fused PET.

3. NODULE SEGMENTATION CHALLENGES

Several complications were met by segmentation techniques of lung nodules, such as ability to segment the nodules, automation level and robustness. We will discuss these challenges in this section.

3.1. Nodule types:

CT scan values of parenchymal tissues vary in nature from soft tissues; the segmentation of large solid nodules is not a complex process, whereas small nodules segmentation-based study, where nodules are attached to the vessels (juxta vascular), nodules attached to parenchymal wall and diaphragm (juxta pleural and ground-glass opacity nodules are performed with higher complexity. Small-nodule detection plays a vital role in early lung cancer detection [62]. Accurate segmentation of small nodules is needed to assess malignancy of lesions by measuring their growth rate. The Partial-volume effect (PVE) is the leading technical concern for handling small nodules.

Juxta vascular nodules refer to nodules that are attached to blood vessels. One common approach for handling juxta vascular cases is morphological filtering [63-68]. The partial volume analysis using segmentation (PVM) approach in order to explore the segmentation of (VOI) into the parenchymal cell [67]. Kuhnigk yma area and partial volume region. Kostis [63] proposed



isometric resampling of volumetric images to mitigate PVE. Because the portion of nodules attached to vessels/airways is tiny for the total extent of the 3D nodule surface, primary MOs, such as erosion, dilation, and opening, are often most effective for most juxta vascular cases [69,70].

Juxta pleural nodules refer to cases that are attached to the parenchymal wall or the diaphragm. Pleural surface removal (PSR) is the most widely adapted method for juxta pleural cases [71-77]. PSR can be addressed globally or locally. The universal protocol segmented the entire lung from the CT image and use the result as a negative mask to avoid non-target wall regions. RAGF proposed by Okada [75] yields an ellipsoidal approximation of the lesion boundary. The local PSR methods remove the pleural surface within a VOI [78-81]. Primary constraint mean shift [82, 83] robust nodule core detection by centricity transformation [84] and the maximum curvature point [85] are examples of robust methods addressing this task.

The ground-glass opacity (GGO) nodules are the sub-solid CT values that are lower than typical solid nodules. Based on reliable components, they are categorized into non-solid/pure and partially solid/mixed. A minor non-solid GGO representing AAH, or BAC can slowly evolve into an invasive lung adenocarcinoma over ten years [86].

3.2. Automation:

Lung nodule segmentation objects as a sub-component of the overall system in CADx. In application context, the projected segmentation methods includes two types: automatic and semi-automatic. The intuitive approach takes a CT scan image as an input and segments the nodules present in the image without specifying the location by users [87-91]. Early methods with grey-scale MM filtering [92] and with automatic locally adaptive thresholding [93] addressed simultaneous segmentation of nodules in volume. On contrary, the semi-automatic approach is assumed that the location of target nodules is known. In this assumption, the segmentation method takes a VOI as as input and hypothesize that the (VOI) contains the overall extent of a single nodule.

3.3. Robustness:

Some standard estimation techniques were implemented for the nodule segmentation problem. A model fitting process is made strong by considering inlier samples. This principle is used in the RAGF method by fitting an anisotropic Gaussian only with pieces defined by the MS [94]. The robust lung surface modelling provides a polynomial surface to the lung wall only with pieces that lie on the pleural surface [95]. The RAGF method imposes ellipsoidal constraints for handling juxtaposed cases. The detection and rough segmentation of nodules are done via the LoG filtering approach [96,97].

3.4. Validation:

It refers to the experimental procedures that measure the accuracy of segmentation methods to be evaluated. The designing of effective validation procedures is a significant challenge in lung nodule segmentation research. The most accurate benchmark for validation currently available is chest phantom scans, where an artificial device stimulating nodules and other parenchymal structures were imaged by a CT scanner.

4. DIAGNOSIS OF LUNG NODULES

The segmentation of lung nodules was followed by classification of the detected nodules into malignant or benign. The CADx system distinguishes the malignant nodules from benign nodules automatically. Receiver-operating characteristic (ROC) analysis evaluates the performance of the CADx system, [98] as it has a 2-class classification. The area falling under the ROC curve represents the performance index [99].

4.1. Diagnosis of lung nodules based on Growth rate:

The growth of small pulmonary nodules is measured in 2D area [100] and [101] 3D volume to predict the malignancy and monitor the tumor response to therapy [102]. In oncology, there exist standard protocols for measuring tumor sizes, such as WHO and RECIST.

A large volume of the recent studies has addressed reproducibility/repeatability of lung nodule volumetry by characterizing its bias and variance. Factors that are covered in these studies include nodule characteristics such as size and shape [103-108] appearance types of solidness [109], pulmonary deformation due to cardiovascular motion [110] and inspiration [111] CT reconstruction parameters such as slice thickness, slice intervals, field of views [112-115] algorithm types [116-118] kernel types [119,120] tube current time [121,122] and dosage settings [123] CT scanner vendors [124] segmentation performance due to choices of threshold



parameters in a segmentation algorithm [125,126] segmentation algorithm [127,128] in the same segmentation software packages [129]. Moreover, versions of a segmentation software package [129] intra- and inter observer varies by using a commercial software package for semi-automatic nodule segmentation and volumetry [130-132] levels of observer’s experience and training [133]. Below is an overview of work which was done for measuring the growth rate of the detected lung nodules: Yankelevitz et al., 2000 [134] used HRCT scans to assess the growth rate of small pulmonary nodules. The ROI containing the nodule was identified in each image manually by a radiologist and then resampled to obtain an isotropic space using a 3D linear interpolation, threshold, and was segmented to reconstruct the 3D image volume. Reeves et al. 2006 [68] presented a method for measuring the change in the nodule size from two CT scan images obtained and resampled the ROI of each CT scan [135] Kawata et al., 2000, [33] co registered the pulmonary nodules by using rigid-body and affine registration at two different stages. The curvature features obtained from the nodules segmented by a 3D deformable surface model track the temporal evolution of the nodule [136]. Jirapatnakul et al 2009 [76] presented a nodule growth measurement method called growth analysis from density (GAD) and applied a Gaussian weighting function to the region around the nodule [137]. El-Baz et al.2011 [30] proposed a method for monitoring the development of lung nodules in successive chest LDCT scans of the patients. (Table 2) depicts different methodologies based on growth rate [138].

Table 2. Growth rate-based methods for following up pulmonary nodules.

STUDY	METHOD	REFERENCES
To determine the accuracy of CT scan volumetric measurements of small pulmonary nodules.	Exponential growth model to estimate the doubling time.	[135]
To develop a method for measuring nodule size change from 2 different CT image scans to establish growth rate.	Registration, adaptive threshold and knowledge-based shape matching method is used.	[68]
To measure nodule growth without explicit segmentation.	Growth analysis from density (GAD)method to measure growth rate	[139]
To monitor the development of lung nodules	GR volumetric measurement	[30]

4.2. Diagnosis of Lung Nodules Based on Shape and Appearance:

A great deal of work is published regarding the usefulness of morphologic features to distinguish between malignant and benign pulmonary nodules on CT. Multiple case studies have reported a correlation between different nodule shape characteristics and their pathology. For example, Furuya et al. 1999 [139] scanned parts of 193 pulmonary nodules based on HRCT scans and grouped them into round, lobulated, halo and ragged. (4/23; Table 1).Kawata and colleagues, 2011 [140] designed a CT-based CAD system that classified pulmonary nodules based on a combination of curvature index and the relationship of the nodules to their surrounding features. The CAD system developed by McNitt-Gray used a pattern based classificational approach for the determination of malignancy at pulmonary nodules on HRCT in a series of 31 cases (17 malignant, 14 benign). Shah et al., 2005 [141] designed a CAD system that extracted features from 2 different contours, i.e., solid portions of the nodule and ground-glass components.

Iwano et al., 2005 [142] developed a system to identify the pulmonary nodules detected on HRCT as variations in the shape categories and compared its performance to radiologists. Thus, various approaches are proposed in CADx systems. Database size varied in further studies; CT scans in the databases included screening LDCT, standard diagnostic CT, and HRCT. Reflections on



the development of CADx systems for distinguishing between malignant and benign lung nodules in CT based on shape and appearance features were summarized in Table 3.

Table 3. Classification between malignant and benign nodules.

PURPOSE	METHOD	OBSERVATION	REFERENCES
To classify the shape of pulmonary nodules by computer analysis of HRCT.	LDA with two features	Out of 102 cases, 7 were mismatched cases with circularity (<0.75) and the second moment (0.18) pulmonary nodules were very likely to reveal lung cancer.	[141]
To classify nodules into benign or malignant.	LDA with stepwise feature selection based on nodules shape, size, attenuation and texture.	A correct classification rate of 90.3%	[139]
To discriminate between B and M nodules	LDA with stepwise feature selection based on nodules features.	Nodule features were more effective than surrounding structure features.	[47]
To classify nodules into benign or malignant	Analysis of the spatial distribution of Hounsfield values of nodules.	Sensitivity of 92.3% and specificity of 96.6% used.	[30]

4.3. PET/CT Nodule Diagnosis:

PET/CT fusion performs tumor staging and pulmonary nodule diagnostics. The images obtained from PET show unregulated metabolism that results in higher FDG uptake that permits the detection of malignancy. Reported studies [140-147] used this characteristic to detect malignant *solitary pulmonary nodules* (SPNs) in PET. SPNs are single, spherical, well-circumscribed, radiographic opacity that measures more than 3 cm in diameter. (See Table 4 for more detail).

Table 4. Evaluation of nodule malignancy in PET.

STUDY	SPECIFICITY	SENSITIVITY
Dewan et al.1993 [142]	80%	95%
Halley et al.2005 [147]	89%	94%
Lee et al.2001 [43]	82%	95%
Gupta et al.1996 [143]	88%	93%
Lowe et al.1998 [44]	90%	92%



Table 5 summarizes the evaluation results of nodule malignancy in fused PET/CT systems. The experiments involved in these studies [148-152] show that PET/CT achieve a higher diagnostic power, so PET/CT fusion presents an advancement in lung cancer applications.

Table 5. Evaluation of nodule malignancy in fused PET/CT.

STUDY	DATABASE	ACCURACY/PPV (Positive, productive value)
Keidar et al.	42	89%
Nakamoto et al.	53	87%
Yi et al.	119	93%
Nie et al.	92	95%

5. TREATMENT BY MOLECULAR PROFILING OF TARGET RECEPTORS

Non-small-cell lung cancers (NSCLC) operate variation in the epidermal growth factor receptor (EGFR), threonine-protein kinase BRAF and chromosome recombination of anaplastic lymphoma kinase (ALK) as well as ROS proto-oncogene receptor tyrosine kinase (ROS1). These receptors are validated targets for kinase-inhibitor therapy [153]. Resistance against targeted therapy occurs due to a de novo adaptation of cellular epigenetic and transcription programs. The introduction of new techniques such as high throughput parallel sequencing highlighting mutation based studies, copy number, promoter and methylation, gene and protein expression studies and metabolic activities occurring in tumors [154, 155]. The Cancer Genome Atlas (TCGA) research described different somatic mutations between tumor and matched pair and scanned tumor DNA, RNA and proteome to identify driver genes and reliable biomarkers in NSCLC and SCLC [156]. Molecular profiling of 230 resected lung adenocarcinomas from all major histologic types designate repeated mutation in 18 genes, of which EGFR frequently mutates, and KRAS rarely mutates. Molecular genotyping is common for lung adenocarcinoma and squamous cell lung cancer patients in today’s life. Efficiency of targeted kinase inhibitors in BRAF, EGFR, ALK and ROS1 gene fusions were disclosed. Molecular based profiling of intrinsic genomic instability in tumors has also appeared. Tumor mutational burden predicts immunotherapy response in NSCLC. [157] Targeted inhibitors are remarkable tools for treating NSCLC patients as NSCLC tumors are insensitive to chemotherapy.

5.1. Kinase inhibitor genes in Lung Adenocarcinoma:

5.1.1. EGFR:

The epidermal growth factor receptors like (EGFR, ErbB-1, HER1 in Homo sapiens) is a transmembrane receptor known as tyrosine kinase that regulates cell growth and differentiation in response to the EGF family of extracellular protein ligands. EGFR dependent growth signaling is facilitated through three distinct downstream signaling arms-MAPK, PI3K/ AKT and the JAK-STAT pathway). Currently, multiple EGFR TKIs have been permitted by the FDA. In response to EGFR TKI, a profound response was observed in tumors that harbor activating EGFR exon 19 deletions and EGFR L858R mutation in exon 21. These tumors account for almost 90% of all EGFR mutations in NSCLC [157]. In 15% of NSCLC patients, somatic mutations in EGFR are found.

5.1.2. BRAF:

It is a member of the Raf kinase group that elevates sign transduction protein kinase and regulates molecular division, growth and differentiation. In TCGA analysis, 3–8% of lung carcinomas harbor somatic changes in BRAF V600E modification [158,159]. The primary BRAF mutations are BRAF G469A/V (BRAF magnificence two mutation) and BRAFD594G (BRAF magnificence three mutation). MEK inhibitors and BRAF inhibitors affect the antitumor efficacy approved by FDA [160].

5.1.3. KRAS:

KRAS is a tiny GTPase in the Ras superfamily, which regulates cell motility, increases survival in reaction to trophic and mitogenic stimuli. Growth signaling through GTP bound activated KRAS are referred through 03 incredible downstream signaling arms as RAF/MAPK, PI3K/AKT and RAL-GEF/RAL signaling pathway [161]. AActivating KRAS mutations are



present in approx. 20–30% of sufferers of NSCLC [162]. The top distinguished KRAS mutant shape is KRAS-G12V locked constitutively in an activated GTP sure state. In pre-scientific models, kinase 1, RhoA/ Rho kinase, nuclear export XPO1 inhibitions result in KRAS mutant lung cancer vulnerability [163,164]. Pharmacological inhibition of upstream adapter protein SHP2 renders reaction in opposition to KRAS-G12C variation. This variation of RAS cycle nucleotide behaves in a semi-constitutive manner. [165].

5.1.4. ALK and ROS1 rearrangements:

Oncogenic ALK and ROS1 gene rearrangements occur in 10% of NSCLC patients [166,167]. The degree of overexpression and activation in fusion proteins depends on the nature of its fusion partner. The most common integrational partner of ALK is EML4 and ROS1 is CD74 [168]. Crizotinib, ceritinib, alectinib and brigatinib are the ALK inhibitors for the treatment of NSCLC. Alectinib was preferred for first line ALK TKI treatment, and improved outcomes of ALEX trials are obtained [169]. Based on the structural homology between the ALK and ROS1 kinase domains, crizotinib was approved for ALK-rearranged lung cancers and ROS1 rearranged cancers [170].

5.2. Kinase inhibitor genes in Squamous cell lung cancer:

5.2.1. FGFR1:

In TCGA analysis, 18% of tumor samples of early-stage squamous cell carcinoma have copy number amplification or mutations on FGFR1 [171]. Zhang and colleagues reported specific sensitivity of PDXs against AZD4547, an FGFR inhibitor [172]. Early attempts testing the FGFR inhibitors, like AZD4547, showed a disappointing rate of 8–12% [173-175]. This poor response rate is attributed to a poor correlation between copy number amplification and protein expression and shows co-alterations in PI3K and cell G1/S checkpoint pathways [176].

5.2.1.1. PI3K/AKT pathway:

PI3K/AKT pathway regulates cell biological signal networks and further regulates cell survival, metabolism and proliferation. In TCGA analysis, alterations in PI3K pathway members, such as PTEN and PIK3CA, have a higher frequency in squamous cell lung cancer [177-178]. Moreover, PTEN and PIK3CA mutations correlate with poor prognosis of human patients [179,180]. The function of biomarkers fails due to adaptive upregulation of PD-L1 receptors in tumors which in turn are immunosuppressive [181-182].

5.2.1.2. CDK4/6:

CDK4/6 and its binding partners Cyclin D1-3s control mitotic cell division and is being negatively regulated by p53/p21 tumor suppressors. The cellular translation potential of RB and the negative regulator of E2F is regulated by CDK4/6 [183]. In squamous cell lung cancer, CCND1 and CyclinD1 are amplified in 15% cases and CDK4/6 mutated in 45% cases. Phase 1 multi-cancer studies of abaculi show partial response in 1 (out of 6) patient with squamous cell carcinoma. The patient harbors CDKN2A loss, which is an inhibitor of CDK4. The list of identified oncogenic drivers in NSCLC is expanding. Targeted therapeutic approaches against these alterations are also currently underway [184-187].

6. MULTI OMICS ANALYSIS IN LUNG CANCER

Multi-omics facts show molecular subtypes of cancers having TCGA dataset. Multi-omics evaluation, integrating TCGA facts of RNA expression, DNA methylation, factor mutations, and replica quantity variation, displays a prediction function for miserably affected person outcomes [188]. So, to broaden a classifier for the prediction of lung cancers affected person diagnosis and analyze an affected person risk-established evaluation, a machine-learning-primarily based pipeline for multi-omics in lung cancers has been applied [189]. Data of RNA expression and miRNA expression as entering facts turned into chosen, so that the result may be interpretable, considering that RNA expression is regulated through miRNA by functional duplexes. Firstly, we evolved an SVM that turned into a position to differentiate diagnosis-associated subtypes from the TCGA LUAD. Secondly, we achieved a risk-established pathway evaluation to present us with useful statistics expertise about ability mechanisms associated with the unique subtypes.

7. METHODS

7.1. TCGA Set:

First, multi-omics LUAD data from the Genomic Data Commons (GDC) TCGA data portal (<https://portal.gdc.cancer.gov>), using



TCGA Assembler 2 (<https://github.com/compgenome365/TCGA-Assembler-2>; with R package (R version 3.5.1) is downloaded. Three hundred eighty-four patients with RNA sequencing and miRNA sequencing data were provided into a multi-omics dataset which was the last step. Patients' clinical data are downloaded from the GDC data portal, and 364 patients were available for the following analysis step. Data with zero values was then processed. In the last step, zero values, RNA-seq data, and miRNA-seq data were done against patients, followed by reassembling to complete a multi-omics dataset before feeding it into an autoencoder.

7.2. Clustering to Obtain Inferred Labels from LUAD Multi-Omics Dataset:

An autoencoder with three hidden layers (Five hundred, hundred, five hundred nodes) with Kera's (<https://keras.io>; version 2.2.4) was executed and L1 and L2 regulation weights were set at 0.001 and 0.0001 for hyperparameter settings. The learning rate is being placed at 0.01, with a 1e-6 decay rate, and epochs of 150, with a 0.5 dropout rate. Stochastic gradient descent (SGD) used to enhance the survival-associated features and to obtain inferred labels; the elbow method [190] is being performed in the clustering range of 1 to 10. Based on the result predicted by the elbow method, further analysis was performed, and the optimal number of clusters, using the Silhouette index [191] and Calinski–Harabasz [192] criterion, were visualized. In the (4/23; Table 1).last step, K-means clustering is performed using the K, and the result was visualized using t-Distributed Stochastic Neighbor Embedding (t-SNE) [193].

7.3. Kaplan–Meier Analysis:

For the Kaplan–Meier analysis and for evaluating the prognostic significance of LUAD patients, inferred labels are gotten by clustering. Survival analysis was performed with the aid of the R survival package, and the survival curve drawn using the R survivor package.

7.4. ANOVA Feature Ranking of miRNA and RNA Expression to Develop SVM Classifier and LUAD Prognosis-Dependent Classifiers:

The multi-omics data required in drawing a Kaplan–Meier plot was divided into 60% for a training dataset and 40% for a test dataset. The analysis of Variance (ANOVA) method was used to rank miRNA and RNA subtypes. ANOVA method with the inferred labels performs through the R limma package. [194] Ranked miRNAs from 5 to 20 and ranked RNAs from 5 to 30 are used to develop SVM. A fixed number of miRNA and RNA were then applied to create another three classifiers: K-nearest neighbors (KNN), Random Forest (RF), and Logistic Regression (LR) and to compare the accuracy with SVM.

7.5. Clinical Characterization:

The clinical information that was being gathered from Kaplan–Meier analysis aids in differentiation between two distinct populations clustered by the K-means algorithm.

7.6. Somatic Mutation Analysis (SNPs and Small Indels):

LUAD somatic mutation data are downloaded from the University of California, Santa Cruz (UCSC) Xena server (<https://xenabrowser.net/datapages/>) and studied for mutations that was occurring in each patient.

7.7. Copy Number Analysis:

LUAD copy number dataset altered by Log2 was taken from the UCSC Xena server (<https://xenabrowser.net/datapages/>). Copy number variant with the Affymetrix SNP 6.0 platform was chosen and assembled by hg38.

7.8. Pathway Analysis Enrichment in the Poor Prognosis Subtype:

Gene set enrichment analysis (GSEA), Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, and Gene Ontology (GO) analysis accomplishes by using DESeq2, and tidyquant packages in R, to compute enriched pathways in the poor survival subtype.

7.9. Identification of the Novel Genes Associated with LUAD Patient Survival:

Expression analysis was carried out between 2 subtypes, using the R DESeq2 package. Then top 25 statistically significant RNAs between the subtypes obtained f described patient survival. The OncoLnc web server helped to draw Kaplan–Meier and get values from each survey (<http://www.oncolnc.org>) [193]. Then, 25% from the high-expression subgroup (n = 123) and 25% from the low-expression subgroup (n = 123) were used to figure prognosis.

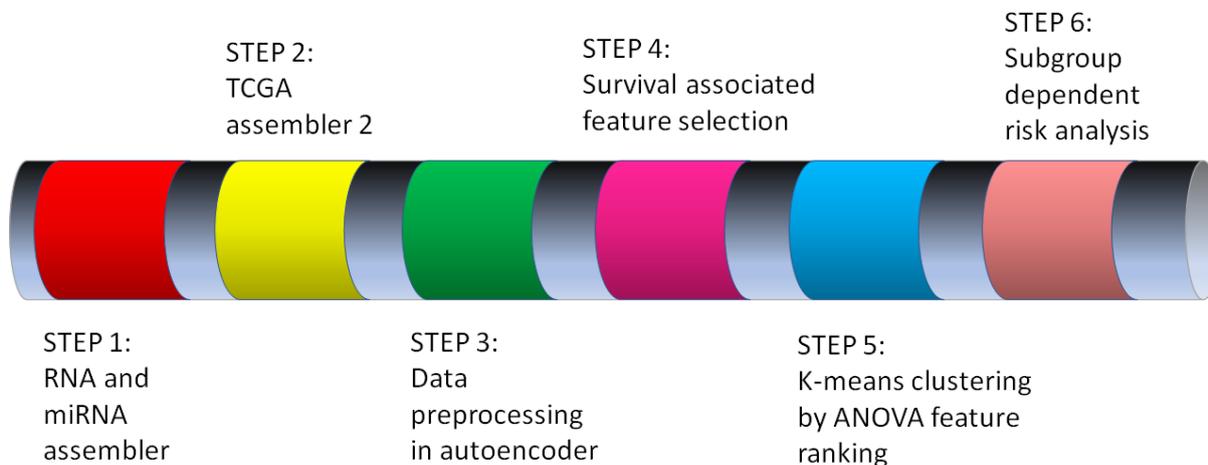


Figure 2. The steps which were followed in a multi-omics setup. In these 6 steps are performed.

8. ROLE OF SIMULATIONS FOR LUNG CANCER WITH RARE EGFR MUTATION

Lung cancer with EGFR mutations is responsible for 10–20% and 40–50% of NSCLC cases [195]. During frame deletions around the LREA motif of exon 19 and exon 21 L858R point mutations are hotspots of EGFR mutations, accounting for ~80–90% of EGFR mutations [196]. All mutations occurring in the EGFR tyrosine kinase domain promote the active conformation of EGFR protein and its oncogenic pathways [197–199]. Multiple EGFR tyrosine kinase inhibitors (EGFR-TKIs) are approved and used in routine cancer clinics to inhibit hyperactive EGFR signaling [200–205] based on a positive relationship between EGFR mutations and sensitivity to EGFR-TKIs has been well-established. EGFR exon 20 transformations consist of more than 50 types and account for 4–10% of all EGFR mutations [206–208]. The wide variety of mutations and presence of many single mutations prevents the portrayal of presently known mutants. Due to the use of NGS tests in lung cancer, the number of novel EGFR mutations are increasing. Hence, a rapid and robust method to precisely guess the sensitivity of EGFR rare mutants is obligatory for solving the problem of NSCLC patients with rare EGFR mutations.

Recently, computational structural modelling and molecular dynamics (MD) simulations have helped us clarify the activation mechanism of EGFR at the atomic level [209–211]. In addition, the sensitivity prediction of EGFR mutants to EGFR tyrosine kinase inhibitors was accomplished for various EGFR mutations using binding free energy calculated with MD simulation [212–213] and fitness scores calculated by molecular docking simulation [214]. Supercomputer-based binding energy is calculated using MD simulation [215–216] and played essential part in the secondary ALK and RET mutants, which is present during therapy using TKIs [217–218]. Based on this, it was observed that a supercomputer-based model would predict the sensitivity of rare EGFR mutants to EGFR-TKIs.

8.1. High Diversity of Rare EGFR Mutations in NSCLC:

To obtain relevant information about individual EGFR mutations and to explore their diversity, we crosschecked with the Japan's Lung Cancer Genomic Screening Project for Individualized Medicine in (LCSCRUM), a prospective nationwide lung cancer clinical and genomic research network [219]. NSCLC cases where the major EGFR mutations (exon 19 deletions, L858R, G719X, or L861Q) were not detected by routine clinical testing underwent NGS evaluation for possible somatic alterations using a panel of cancer-related genes. Within LC-SCRUM-Japan, 3,779 NSCLC patients were registered from February, 2013 to March 2017. Out of these patients, 201 from October, 2013 to June, 2014 were grouped as first cohort and patients 1,963 from March, 2015 to March, 2017 grouped into second cohort were subjected to next gen. sequencing NGS. The NGS research showed that the major EGFR mutations with 19 exonic deletions, and G719X, or L861Q and L858R, were detected in 2.7% (53/ 1,963) of such patients in the second cohort, indicating false-negative results of daily based clinical tests. The rare EGFR mutations were detected in 108 (5.5%) patients. The frequency of rare EGFR mutations was higher than ROS1 fusions (3.6%) and RET fusions (2.9%). However, the frequency of rare EGFR mutations comprised approximately one-third of KRAS mutation frequency (15.3%). Rare EGFR mutations are noticed in both non squamous NSCLC and squamous cell carcinoma. These data highlighted that rare EGFR



mutations account for a significant proportion of NSCLC cases. Most mutations are found in the region encoding the tyrosine kinase domain, in exonic region of 18–21, while some mutations are at external region to the tyrosine kinase domain, particularly in the exonic region (6, 7, 8, 12, 15, and 17). The most uncommonly encountered EGFR mutations were EGFR exon 20 insertion mutations. Out of which the total of 113 NSCLC cases with rare EGFR mutations, the five cases in the first cohort and 108 cases in the second cohort, 52 (46.0%) harbored EGFR exon 20 insertion mutations, indicated that the mutations comprised about a half of rare EGFR mutations. Of the identified 73 rare EGFR mutations, 68 types were observed in only 1-2 cases. These findings indicate a high diversity of EGFR mutations in NSCLC.

8.2. Low Chance of EGFR-TKI Therapy for Rare EGFR Mutation Cases:

Clinical data are available for 53 patients with EGFR major mutations, 47 patients with exon 20 insertion mutations and 61 other with rare mutations in NSCLC. The characteristics highlighted in the patients of NSCLC with exon 20 insertion mutations are similar to patients with significant mutations. However, the nature of NSCLC patients with other rare mutations was slightly different. The majority features of rare EGFR mutations were mutually exclusive with other oncogene mutations, however, co-occurrence with ERBB2 amplification, PIK3CA mutation, KRAS mutation, and MET amplification is observed in a small subset of patients. Also, out of the total 113 patients with uncommon EGFR mutations, 82 patients, including 33 cases with the mutation at exon 20, were reported for information on chemotherapy after gene testing. However, out of those remaining 82 patients, 77 (93.9%) undergone cytotoxic chemotherapy as the primary treatment, while only five patients (6.1%) undergo EGFR-TKIs as other primary treatment approach. About 32 patients (97.0%) with EGFR exon 20 insertion mutations are treated with cytotoxic chemotherapy as first phase of treatment. The overall response rate to EGFR-TKIs in the patients with rare EGFR mutations was only 17.4%. Notably, one of these four cases, who responded to afatinib, carried NSCLC with an exon 20 insertion mutation [220]. These resulted in the validation of overall response to first- and second-generation EGFR-TKIs for NSCLC cases with uncommon EGFR mutations is low, there is a responsive sub-group among them.

8.3. Prediction of sensitivity of EGFR mutants:

MD Simulations and Molecular Docking research are expecting the sensitivity of EGFR mutants to TKIs [221-223]. First, to evaluate the usefulness of the molecular docking approach with our model, we calculated the fitness (docking) rating of TKIs to EGFR mutants by using rDock [224]. One of the maxima broadly used molecular docking applications for structure-primarily based digital screening. Next step was to assess several MD simulation strategies with the model, using Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) and Molecular Mechanics Generalized Born Surface Area (MM-GBSA) with the scripted python MMPBSA.py module [225] in Amber Tools was being used, it was as aid in enquiring the sensitivity of EGFR mutants to TKIs [226]. Besides, EGFR exon 20 insertion mutations to EGFR-TKIs are used for in-silico prediction techniques based on MD simulation.

9. CONCLUSION

Lung cancer is one of the most widespread health issues worldwide whose mortality rate is high. The main risk factor of lung cancer is smoking. Prevention can be by avoiding risk factors, smoking and air pollution. Better outcomes are possible only in the case of early detection. Hence designing an efficient CAD system for lung cancer is very important as early diagnosis can improve the effectiveness of treatment. In this paper, we discussed the current approaches, strengths, limitations developed in each stage of the CAD system. It also explores the kinase inhibitor genes in lung adenocarcinoma and squamous cell lung cancer. Recently, multi-omics analysis has emerged to allow the classification of groups of patients based on prognosis. Here we have addressed the combination of RNA expression, miRNA expression, and clinical information for the development of SVM to predict patient survival in lung cancer. The paper reveals diversity and driver roles of rare EGFR mutations in a large perspective. The frequency of each rare EGFR mutation reports less than 0.1%.

Author's Contribution

PP, PK, NK, KSM, SM, SN, SR, NN, RV did intensive research on various 'lung cancer' topics developed; PP, SG, PK contributed in writing the manuscript; Scientist MN designed and supervised the present review article and assisted in writing the paper.



REFERENCES

1. Lung Carcinoma: Tumors of the Lungs .Merck Manual Professional Edition. online edition. Retrieved 15th August 2007.
2. Non-small cell lung cancer Treatment-Patient Version (PDQ®).NCL. May 12,2015. retrieved 5 March 2016.
3. Stewart, edited by Bernard W, Wild, Christopher P. World cancer report2014. Lyon: IARC Press pp.350-352.ISBN 9789283204299.
4. World Cancer Report 2014.World Health Organization. 2024. pp. Chapter 5.1.ISBN 9284304298.
5. National Cancer Registry Report. Malaysia Cancer Statistics- Data and Figure.2007.
6. Thun MJ, Hannan LM, Adams CambellLL ,et al. Lung cancer occurrence in never smokers: an analysis of 13 cohorts and 22 cancer registry studies.PLoSMed.2008;5(9):1185.
7. O'Reilly KM, Mclaughlin AM, Beckett WS, et al. Asbestos related lung disease American Family Physician.2007; 75(5):683-688.
8. Lung Cancer-patient Version. NCI .Retrieved 5 March 2016.
9. Hom L, Lovly CM, Johnson DH (2015).Chapter 107.Neoplasm of Lung .In Kasper DL ;Hauser, SL, Jameson JL, Fauci AS, Longo DL, Loscalzo J. Harrison's Principles of Internal Medicine(19th ed.) McGraw- Hill.ISBN 978-0-07-180216-1.
10. Lu C, Onn A, Vaporciyan AA, et al.(2010).78:Cancer of Lung. Holland- Frei Cancer Medicine(8th ed.).People's Medical Publishing House. ISBN 978-1-60795-014-1.
11. Lung Carcinoma:Tumors of the Lungs. Merck Manual Professional Edition.online edition. Retrieved 15th August 2007.
12. Chapman S, Robinson G, Streadling J, et al (2009).Chapter 31.Oxford Handbook of Respiratory Medicine(2nd ed.) Oxford University Press .ISBN 978-0-19-954516-2.
13. World Cancer Report 2014.World Health Organization. 2014.pp. Chapter 1.1.ISBN 9283204298.
14. Jemal A, Tiwari RC, Murray T, et al. Cancer Statistics 2004.CA:A Cancer Journal for Clinicians.2004;54(1):8-29.
15. Hom L, Lovly CM, Johnson DH (2015). Chapter 107. Neoplasm of Lung. In Kasper DL; Hauser, SL, Jameson JL, Fauci AS, Longo DL, Loscalzo J. Harrison's Principles of Internal Medicine (19th ed.) McGraw- Hill.ISBN 978-0-07-180216-1.
16. Brown KM, Keats JJ, Sekulie A, et al. (2010). Chapter 8 Holland –Frei Cancer Medicine (8th ed.). People's Medical Publishing House USA.ISBN 978-1-60795-014-1.
17. Cooper WA, Lam DLC, OTooleSA.Molecular Biology of lung cancer (PDF). J ThoracDis.2013;5(Suppl 5):S479-90.
18. Tobias J, Hochhauser D (2012). Chapter 12.Cancer and its Management (6th ed.) Wiley-Blackwell.p. 200.ISBN 978-1405-170154.
19. Herbst RS, Heymatch JV, Lippman SM. Lung cancer. NewEngl JMed. 2008; 359(13): 1367-1380.
20. Kumar V, Abbas AK, Aster JC (2013). Chapter 5.Robbins Basic Pathology(9th ed.)Elsevier Saunders.p.212.ISBN 978-1-4377-178-5.
21. Jakopovic, M, Thomas A, Balasubramaniam S .Targeting the epigenomic in lung cancer expanding approaches to epigenetic therapy(PDF).Frontiers in oncology. 2013;3(26).
22. Herbst RS, Heymatch JV, Lippman SM. Lung cancer. NewEngl JMed. 2008; 359(13): 1367-1380.
23. Kumar V, Abbas AK, Aster JC (2013).Chapter 5.Robbins Basic Pathology(9th ed.)Elsevier Saunders.p.212.ISBN 978-1-4377-178-5.
24. Herbst RS, Heymatch JV, Lippman SM. Lung cancer. NewEngl JMed. 2008; 359(13): 1367-1380.
25. Hu S, Hoffman EA, and Reinhardt JM. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. IEEE Transactions on Medical Imaging. 2001; vol. 20(6):490–498,.
26. Ukil S and Reinhardt JM. Anatomy-guided lung lobe segmentation in X-ray CT images. IEEE Transactions on Medical Imaging. 2009; 28 (2); 202–214.
27. Van Rikxoort EM, De Hoop B, Van De Vorst S, Prokop M, and Van Ginneken B. Automatic segmentation of pulmonary segments from volumetric chest CT scans. IEEE Transactions on Medical Imaging. 2009; 28(4): 621–630,
28. Gao Q, Wang S, Zhao D, and Liu J. Accurate lung segmentation for X-ray CT images,” in Proceedings of the 3rd International Conference on Natural Computation (ICNC '07). 2007: 2: 275– 279,



29. Hua P, Song Q, Sonka M, Hoffman EA, and Reinhardt JM. Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm, in Proceedings of the International Symposium on Biomedical Imaging (ISBI' 11). 2011; 2072–2075.
30. El-Baz A, Gimel'farb G, and Falk R. A novel three-dimensional framework for automatic lung segmentation from low dose computed tomography images, in Lung Imaging and Computer Aided Diagnosis, A. El-Baz and J. Suri, Eds., chapter 1. 2011; 1–15.
31. El-Baz A, Gimel'farb G, Falk R, Abou El-Ghar M, Holland T and Shaffer T. A new stochastic framework for accurate lung segmentation, in Proceedings of the International Conference on Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '08). 2008; 322–330.
32. El-Baz A, Gimel'farb G, Falk R, Holland T and Shaffer T. A framework for unsupervised segmentation of lung tissues from low dose computed tomography images,” in Proceedings of the British Machine Vision Conference. 2008: 855–865. University of Leeds, Leeds, UK, September.
33. K. Kanazawa, Y. Kawata, N. Niki et al. Computer-aided diagnosis for pulmonary nodules based on helical CT images,” Computerized Medical Imaging and Graphics. 1998; 22(2):157–167.
34. Trivedi MM and Bezdek JC. Low-level segmentation of aerial images with fuzzy clustering,” IEEE Transactions on Systems, Man and Cybernetics. 1986; 16 (4): 589–598.
35. Mekada Y, Kusanagi T, Hayase Y. et al.. Detection of small nodules from 3D chest X-ray CT images based on shape features, in Proceedings of the Computer Assisted Radiology and Surgery (CARS). 2003; 1256: 971–976,
36. Paik DS, Beaulieu CF, Rubin GD et al. Surface normal overlap: a computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT. IEEE Transactions on Medical Imaging. 2004; 23(6):661–675.
37. Li Q, Li F, Suzuki K. et al.. Computer-aided diagnosis in thoracic CT, Seminars in Ultrasound, CT and MRI. 2005; 26(5): 357–363.
38. Li Q. Recent progress in computer-aided diagnosis of lung nodules on thin-section CT,” Computerized Medical Imaging and Graphics. 2007; 31(4-5): 248–257.
39. Goldin JG, Brown MS, and Petkovska I. Computer aided diagnosis in lung nodule assessment. Journal of Thoracic Imaging. 2008; 23(2):97–104.
40. Sluimer I, Schilham A, Prokop M and Van Ginneken B. Computer analysis of computed tomography scans of the lung: a survey. IEEE Transactions on Medical Imaging. 2006; 25(4):385–405.
41. Jaffe CC. Measures of response: RECIST, WHO, and new alternatives. Journal of Clinical Oncology. 2006; 24(20):3245–3251.
42. Gavrielides MA, Kinnard LM, Myers KJ and Petrick N. Noncalcified lung nodules: volumetric assessment with thoracic CT. Radiology. 2009; 251(1): 26–37.
43. Lee Y, Hara T, Fujita H, Itoh S and Ishigaki T. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. IEEE Transactions on Medical Imaging. 2001; 20(7):595–604.
44. Brown MS, McNitt-Gray MF, Goldin JG, Suh RD, Sayre JW, and Aberle DR. Patient-specific models for lung nodule detection and surveillance in CT images. IEEE Transactions on Medical Imaging. 2001; 20(12): 1242–1250.
45. Armato SG, Li F, Giger ML, MacMahon H, Sone S and Doi K. Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program, Radiology. 2002; 225(3):685–692.
46. Suzuki K, Shiraishi J, Abe H, MacMahon H, and Doi K. False-positive reduction in computer-aided diagnostic scheme for detecting nodules in chest radiographs by means of massive training artificial neural network. Academic Radiology. 2005; 12(2):191–201.
47. Kawata Y, Niki N, Ohmatsu H et al. Classification of pulmonary nodules in thin-section CT images based on shape characterization, in Proceedings of the International Conference on Image Processing (ICIP '97). 1997; 3 (2) (of 3): 528–530.
48. Gonzalez RC and Woods RE. Digital Image Processing, Prentice Hall, 3rd edition, 2007.
49. Gonzalez RC and Woods RE. Digital Image Processing, Prentice Hall, 3rd edition, 2007.
50. Kass M, Witkin A, and Terzopoulos D. Snakes: active contour models. International Journal of Computer Vision. 1988;



- 1(4):321–331.
51. Caselles V, Kimmel R, and Sapiro G. Geodesic Active Contours,” *International Journal of Computer Vision*. 1997; 22(1): 61–79.
 52. Chan TF and Vese LA. Active contours without edges. *IEEE Transactions on Image Processing*. 2001; 10(2):266–277.
 53. Amini AA, Weymouth TE and Jain RC. Using dynamic programming for solving variational problems in vision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990; 12(9):855–867.
 54. Lee Y, Hara T, Fujita H, Itoh S and Ishigaki T. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Transactions on Medical Imaging*. 2001; 20(7):595–604
 55. Duda RO, Hart PE and Stork DG, *Pattern Classification*, Wiley Interscience, New Jersey, NJ, USA, 2nd edition, 2001
 56. Duda, RO, Hart PE and. Stork DG, *Pattern Classification*, Wiley Interscience, New Jersey, NJ, USA, 2nd edition, 2001
 57. Comaniciu D and Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24 (5): 603–619.
 58. Boykov Y, Veksler O, and Zabih R, Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2001; 23(11): 1222–1239.
 59. Gonzalez RC and Woods RE, *Digital Image Processing*, Prentice Hall, 3rd edition, 2007.
 60. Steenbakkens RJHM, Duppen JC, Fitton I et al. Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis. *International Journal of Radiation Oncology Biology Physics*. 2006; 64 (2): 435–448.
 61. Fox JL, Rengan R, OMeara W et al. Does registration of PET and planning CT images decrease interobserver and intraobserver variation in delineating tumor volumes for non-small-cell lung cancer? *International Journal of Radiation Oncology Biology Physics*. 2005; 62(1): 0–75.
 62. Henschke CI, McCauley DI, Yankelevitz DF et al. Early lung cancer action project: overall design and findings from baseline screening. *The Lancet*. 1999; 354 (9173): 99–105.
 63. Kostis WJ, Reeves AP, Yankelevitz DF, and Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Transactions on Medical Imaging*. 2003; 22(10):1259–1274.
 64. Fetita CI, Prteux F, Beigelman-Aubry C and Grenier P. 3D automated lung nodule segmentation in HRCT in *Proceedings of the International Conference Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '03)*. 2003; 2878: 626–634.
 65. Zhang L, Fang M, Naidich DP and Novak CL. Consistent interactive segmentation of pulmonary ground glass nodules identified in CT studies in *Medical Imaging: Imaging Processing, Proceedings of SPIE*. 2004; 1709–1719.
 66. Zhao B, Reeves AP, Yankelevitz DF and Henschke CI. Three-dimensional multicriterion automatic segmentation of pulmonary nodules of helical computed tomography images. *Optical Engineering*. 1999; 38(8): 1340–1347.
 67. Kuhnigk JM, Dicken V, Bornemann L et al. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans,” *IEEE Transactions on Medical Imaging*. 2006; 25(4): 417–434.
 68. Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B and Kostis WJ. On measuring the change in size of pulmonary nodules. *IEEE Transactions on Medical Imaging*. 2006; 25(4): 435–450.
 69. Fetita CI, Prteux F, Beigelman-Aubry C and Grenier P. 3D automated lung nodule segmentation in HRCT in *Proceedings of the International Conference Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '03)*. 2003; 2878: 626–634,
 70. Zhao B, Reeves AP, Yankelevitz DF and Henschke CI. Three-dimensional multicriterion automatic segmentation of pulmonary nodules of helical computed tomography images. *Optical Engineering*. 1999; 38 (8): 1340–1347.
 71. Kostis WJ, Reeves AP, Yankelevitz DF and Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Transactions on Medical Imaging*. 2003; 22(10):1259–1274.
 72. Fetita CI, Prteux F, Beigelman-Aubry C and Grenier P. 3D automated lung nodule segmentation in HRCT,” in *Proceedings of the International Conference Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '03)*. 2003; 2878: 626–634.
 73. Kuhnigk JM, Dicken V, Bornemann L et al. Morphological segmentation and partial volume analysis for volumetry of



- solid pulmonary lesions in thoracic CT scans. *IEEE Transactions on Medical Imaging*. 2006; 25(4): 417–434.
74. Shen H, Goebel B and Odry B. A new algorithm for local surface smoothing with application to chest wall nodule segmentation in lung CT data in *Medical Imaging: Imaging Processing*. 2004; 5370; 1519–1526.
 75. Okada K, Ramesh V, Krishnan A, Singh M and Akdemir U. Robust pulmonary nodule segmentation in CT: improving performance for juxtapleural cases in *Proceedings of the International Conference on Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '05)*. 2005; 8:781–789.
 76. Jirapatnakul YD, Mulman YD, Reeves AP, Yankelevitz DF and Henschke CI. Segmentation of juxtapleural pulmonary nodules using a robust surface estimate. *International Journal of Biomedical Imaging*. 2011; 14
 77. Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B and Kostis WJ. On measuring the change in size of pulmonary nodules,” *IEEE Transactions on Medical Imaging*, vol. 25, no. 4, pp. 435–450, 2006.
 78. Kostis WJ, Reeves AP, Yankelevitz DF and Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images,” *IEEE Transactions on Medical Imaging*. 2003; 22(10): 1259–1274.
 79. Okada K, Ramesh V, Krishnan A, Singh M and Akdemir U. Robust pulmonary nodule segmentation in CT: improving performance for juxtapleural cases in *Proceedings of the International Conference on Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '05)*. 2005; 8: 781–789.
 80. Jirapatnakul YD, Mulman YD, Reeves AP, Yankelevitz DF and Henschke CI. Segmentation of juxtapleural pulmonary nodules using a robust surface estimate. *International Journal of Biomedical Imaging*. 2011; Article ID 632195, 14 pages.
 81. Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B and Kostis WJ. On measuring the change in size of pulmonary nodules. *IEEE Transactions on Medical Imaging*. 2006; 25(4): 435–450.
 82. Okada K, Singh M, and Ramesh V. Prior-constrained scale space mean shift,” in *Proceedings of the British Machine Vision Conference*. 2006; 829–838,
 83. T Kubota, Jerebko AK, Dewan M, Salganicoff M and Krishnan A. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Medical Image Analysis*. 2011; 15(1):133–154.
 84. Farag AA, Abdelmunim H, Graham J et al. Variational approach for segmentation of lung nodules in *Proceedings of the IEEE International Conference on Image Processing (ICIP '11)*. 2011; 2157–2160,
 85. Min JH, Lee HY, Lee KS et al. Stepwise evolution from a focal pure pulmonary ground-glass opacity nodule into an invasive lung adenocarcinoma: an observation for more than 10 years. *Lung Cancer*. 2010; 69(1):123–126.
 86. Fetita CI, Prteux F, Beigelman-Aubry C and Grenier P. 3D automated lung nodule segmentation in HRCT in *Proceedings of the International Conference Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '03)*. 2003; 2878: 626–634.
 87. Mullally W, Betke M, Wang J and Ko JP. Segmentation of nodules on chest computed tomography for growth assessment. *Medical Physics*. 2004; 31(4):839–848.
 88. Zhou J, Chang S, Metaxas DN, Zhao B, Ginsberg MS and Schwartz LH. An automatic method for ground-glass opacity nodule detection and segmentation from CT studies in *Proceedings of the 29th IEEE Annual International Conference of Medicine and Biology Society (EMBS '06)*. 2006; 1:3062–3065.
 89. Zhou J, Chang S, Metaxas DN, Zhao B, Schwartz LH and Ginsberg MS. Automatic detection and segmentation of ground-glass opacity nodules in *Proceedings of the International Conference on Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '06)*. 2006; 9:784–791.
 90. Tao Y, Lu L, Dewan M et al. Multi-level ground glass nodule detection and segmentation in CT lung images in *Proceedings of the International Conference Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '09)*. 2009; 1: 715– 723.
 91. Fetita CI, Prteux F, Beigelman-Aubry C and Grenier P. 3D automated lung nodule segmentation in HRCT,” in *Proceedings of the International Conference Medical Imaging Computing and Computer-Assisted Intervention (MICCAI '03)*. 2003; 2878P: 626–634.
 92. Mullally W, Betke M, Wang J, and Ko JP. Segmentation of nodules on chest computed tomography for growth assessment. *Medical Physics*. 2004; 31(4): 839–848.
 93. Okada K, Comaniciu D and Krishnan A. Robust anisotropic Gaussian fitting for volumetric characterization of



- pulmonary nodules in multislice CT. *IEEE Transactions on Medical Imaging*. 2005; 24(2): 409–423.
94. Jirapatnakul YD, Mulman YD, Reeves AP, Yankelevitz DF and Henschke CI. Segmentation of juxtapleural pulmonary nodules using a robust surface estimate. *International Journal of Biomedical Imaging*. 2011; Article ID 632195, 14 pages.
95. Jirapatnakul AC, Fotin SV, Reeves AP, Biancardi AM, Yankelevitz DF and Henschke CI. Automated nodule location and size estimation using a multi-scale Laplacian of Gaussian filtering approach in Proceedings of the 29th IEEE Annual International Conference of Medicine and Biology Society (EMBS '09).2009; 1028–1031.
96. Diciotti S, Lombardo S, Coppini G, Grassi L, Falchini M and Mascacchi M. The LOG characteristic scale: a consistent measurement of lung nodule size in CT imaging. *IEEE Transactions on Medical Imaging*. 2010; 29(2): 397–409.
97. Metz CE. ROC methodology in radiologic imaging. *Investigative Radiology*. 1986; 21(9); 720–733.
98. Hanley JA and McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*.1983;148(3):839–843.
99. Yankelevitz DF, Gupta R, Zhao B and Henschke CI. Small pulmonary nodules: evaluation with repeat CT—preliminary experience. *Radiology*.1999; 212(2): 561–566.
100. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B and Henschke CI. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology*. 2000; 217 (1): 251–256.
101. Jaffe CC. Measures of response: RECIST, WHO, and new alternatives,” *Journal of Clinical Oncology*.2006; 24(20): 3245–3251.
102. Kostis WJ, Yankelevitz DF, Reeves AP, Fluture SC and Henschke CI. Small pulmonary nodules, reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. *Radiology*. 2004; 231(2):446–452.
103. Goo JM, Tongdee T, Tongdee R, Yeo K, Hildebolt CF and Bae KT. Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy. *Radiology*. 2005; 235(3): 850–856.
104. Gietema HA, Schaefer-Prokop CM, Mali WPTM, Groenewegen G and Prokop M. Pulmonary nodules: interscan variability of semiautomated volume measurements with multisection CT - Influence of inspiration level, nodule size, and segmentation performance. *Radiology*. 2007; 245(3): 889–894.
105. Das M, Ley-Zaporozhan J, Gietema HA et al. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. *European Radiology*. 2007; 17(8): 1979–1984.
106. Tao P, Griess F, Lvov Y et al. Characterization of small nodules by automatic segmentation of X-ray computed tomography images. *Journal of Computer Assisted Tomography*. 2004; 28(3): 372–377.
107. Petrou M, Quint LE, Nan B, and Baker LH. Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. *American Journal of Roentgenology*.2007; 188(2):306–312.
108. Park CM, Goo JM, Lee HJ, Kim KG, Kang MJ and Shin YH. Persistent pure ground-glass nodules in the lung: interscan variability of semiautomated volume and attenuation measurements. *American Journal of Roentgenology*.2010; 195(6): W408–W414.
109. Boll DT, Gilkeson RC, Fleiter TR, Blackham KA, Duerk JL and Lewin JS. Volumetric assessment of pulmonary nodules with ECG-gated MDCT. *American Journal of Roentgenology*. 2004; 183(5):1217–1223,
110. Gietema HA, Schaefer-Prokop CM, Mali WPTM, Groenewegen G and Prokop M. Pulmonary nodules: interscan variability of semiautomated volume measurements with multisection CT - Influence of inspiration level, nodule size, and segmentation performance. *Radiology*. 2007; 245(3):889–894.
111. Way TW, Chan HP, Goodsitt MM et al. Effect of CT scanning parameters on volumetric measurements of pulmonary nodules by 3D active contour segmentation: a phantom study. *Physics in Medicine and Biology*.2008; 53(5):1295–1312.
112. Goo JM, Tongdee T, Tongdee R, Yeo K, Hildebolt CF, and Bae KT. Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy. *Radiology*. 2005; 235(3): 850–856.
113. Ravenel JG, Leue WM, Nietert PJ, Miller JV, Taylor KK and Silvestri GA. Pulmonary nodule volume: effects of reconstruction parameters on automated measurements—a phantom study. *Radiology*.2008; 247(2): 400–408.
114. Tao P, Griess F, Lvov Y et al. Characterization of small nodules by automatic segmentation of X-ray computed



- tomography images. *Journal of Computer Assisted Tomography*. 2004; 28(3):372–377.
115. Petrou M, Quint LE, Nan B, and Baker LH. Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. *American Journal of Roentgenology*. 2007; 188(2):306–312.
116. Honda O, Sumikawa H, Johkoh T et al. Computer-assisted lung nodule volumetry from multi-detector row CT: influence of image reconstruction parameters, *European Journal of Radiology*. 2007; 62(1):106–113.
117. Ko JP, Rusinek H, Jacobs EL et al. Small pulmonary nodules: volume measurement at chest CT—phantom study. *Radiology*. 2003; 228(3):864–870.
118. Way TW, Chan HP, Goodsitt MM et al. Effect of CT scanning parameters on volumetric measurements of pulmonary nodules by 3D active contour segmentation: a phantom study. *Physics in Medicine and Biology*. 2008; 53(5):1295–1312.
119. Honda O, Sumikawa H, Johkoh T et al. Computer-assisted lung nodule volumetry from multi-detector row CT: influence of image reconstruction parameters. *European Journal of Radiology*. 2007; 62(1): 106–113.
120. Park CM, Goo JM, Lee HJ, Kim KG, Kang MJ and Shin YH. Persistent pure ground-glass nodules in the lung: interscan variability of semiautomated volume and attenuation measurements. *American Journal of Roentgenology*. 2010; 195(6): W408–W414.
121. Ravenel JG, Leue WM, Nietert PJ, Miller JV, Taylor KK and Silvestri GA. Pulmonary nodule volume: effects of reconstruction parameters on automated measurements—a phantom study. *Radiology*. 2008; 247(2):400–408.
122. Ko JP, Rusinek H, Jacobs EL et al. Small pulmonary nodules: volume measurement at chest CT—phantom study. *Radiology*. 2003; 228(3):864–870.
123. Way TW, Chan HP, Goodsitt MM et al. Effect of CT scanning parameters on volumetric measurements of pulmonary nodules by 3D active contour segmentation: a phantom study. *Physics in Medicine and Biology*. 2008; 53(5):1295–1312.
124. Hein PA, Romano VC, Rogalla P et al. Variability of semiautomated lung nodule volumetry on ultralow-dose ct: comparison with nodule volumetry on standard-dose CT. *Journal of Digital Imaging*. 2010; 23(1):8–17.
125. Das M, Ley-Zaporozhan J, Gietema HA et al. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. *European Radiology*. 2007; 17(8):979–1984.
126. Goo JM, Tongdee T, Tongdee R, Yeo K, Hildebolt CF, and Bae KT. Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy. *Radiology*. 2005; 235(3): 850–856.
127. Lee MC, Wiemker R, Boroczky L et al. Impact of segmentation uncertainties on computer-aided diagnosis of pulmonary nodules,” *International Journal of Computer Assisted Radiology and Surgery*. 2008; 3(6): 551–558.
128. Ko JP, Rusinek H, Jacobs EL et al. Small pulmonary nodules: volume measurement at chest CT—phantom study. *Radiology*. 2003; 228(3): 864–870.
129. Lee MC, Wiemker R, Boroczky L et al. Impact of segmentation uncertainties on computer-aided diagnosis of pulmonary nodules. *International Journal of Computer Assisted Radiology and Surgery*. 2008; 3(6):551–558.
130. Das M, Ley-Zaporozhan J, Gietema HA et al. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. *European Radiology*. 2007; 17(8):1979–1984.
131. Sone S, Tsushima K, Yoshida K, Hamanaka K, Hanaoka T and Kondo R. Pulmonary Nodules. Preliminary Experience with Semiautomated Volumetric Evaluation by CT Stratum. *Academic Radiology*. 2010; 17(7):900–911.
132. Ashraf H, de Hoop B, Shaker SB et al. Lung nodule volumetry: segmentation algorithms within the same software package cannot be used interchangeably,” *European Radiology*. 2010; 20(8):1878–1885.
133. Furuya K, Murayama S, Soeda H et al. New classification of small pulmonary nodules by margin characteristics on high resolution CT. *Acta Radiologica*. 1999; 40(5):496–504.
134. Kawata Y, Niki N, Ohmatsu H et al. Computerized analysis of 3-D pulmonary nodule images in surrounding and internal structure feature spaces, in *Proceedings of IEEE International Conference on Image Processing (ICIP '01)*. 2001; 2: 889–892.
135. McNitt-Gray MF, Hart EM, Wyckoff N, Sayre JW, Goldin JG and Aberle DR. A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results. *Medical Physics*. 1999; 26(6):880–888.



136. Shah SK, McNitt-Gray MF, Rogers SR et al. Computer aided diagnosis of the solitary pulmonary nodule. *Academic Radiology*. 2005; 12(5):570–575.
137. Iwano S, Nakamura T, Kamioka Y, and Ishigaki T. Computer-aided diagnosis: a shape classification of pulmonary nodules imaged by high-resolution CT. *Computerized Medical Imaging and Graphics*. 2005; 29(7):565–570.
138. Dewan NA, Gupta NC, Redepenning LS, Phalen JJ and Frick MP. Diagnostic efficacy of PET-FDG imaging in solitary pulmonary nodules: potential role in evaluation and management. *Chest*. 1993; 104(4): 997–1002.
139. Gupta NC, Maloof J, and Gunkel E. Probability of Malignancy in Solitary Pulmonary Nodules Using Fluorine-18-FDG and PET. *Journal of Nuclear Medicine*. 1996; 37(6): 943–948.
140. Lowe VJ, Fletcher JW, Gobar L et al. Prospective investigation of positron emission tomography in lung nodules. *Journal of Clinical Oncology*. 1998; 16(3):1075–1084.
141. Lee J, Aronchick JM, and Alavi A. Accuracy of F-18 fluorodeoxyglucose positron emission tomography for the evaluation of malignancy in patients presenting with new lung abnormalities: a retrospective review. *Ches*. 2001; 120(6):1791–1797.
142. Herder GJ, Golding RP, Hoekstra OS et al. The performance of 18F-fluorodeoxyglucose positron emission tomography in small solitary pulmonary nodules. *European Journal of Nuclear Medicine and Molecular Imaging*. 2004; 31(9):31–1236.
143. Halley A, Hugentobler A, Icard P et al. Efficiency of 18F-FDG and 99mTc-depreotide SPECT in the diagnosis of malignancy of solitary pulmonary nodules. *European Journal of Nuclear Medicine and Molecular Imaging*. 2005; 32(9):1026–1032.
144. Erasmus JJ, McAdams HP and Connolly JE. Solitary pulmonary nodules: part II. Evaluation of the indeterminate nodule. *Radiographic*. 2000; 20(1): 59–66.
145. Goo JM, Im JG, Do KH et al. Pulmonary tuberculoma evaluated by means of FDG PET: findings in 10 cases. *Radiology*. 2000; 216(1):117-1212.
146. Keidar Z, Haim N, Guralnik L et al. PET/CT using 18F FDG in suspected lung cancer recurrence: diagnostic value and impact on patient management. *Journal of Nuclear Medicine*. 2004; 45(10):1640–1646.
147. Yi CA, Kyung SL, Kim BT et al. Tissue characterization of solitary pulmonary nodule: comparative study between helical dynamic CT and integrated PET/CT. *Journal of Nuclear Medicine*. 2006; 47(3):443–450.
148. Nie Y, Li Q, Li F, Pu Y, Appelbaum D and Doi K. Integrating PET and CT information to improve diagnostic accuracy for lung nodules: a semiautomatic computer-aided method. *Journal of Nuclear Medicine*. 2006; 47(7):1075–1080, 2006.
149. Nakamoto Y, Senda M, Okada T. et al. Software-based fusion of PET and CT images for suspected recurrent lung cancer. *Molecular Imaging and Biology*. 2008; 10 (3): 147–153.
150. Lardinois D, Weder W, Hany TF et al. Staging of non-small cell lung cancer with integrated positron-emission tomography and computed tomography. *The New England Journal of Medicine*. 2003; 348(25):2500–2507.
151. Rotow J, Bivona TG. Understanding and targeting resistance mechanisms in NSCLC. *Nat Rev Cancer* 2017; 17:637–58.
152. Blakely CM, Pazarentzos E, Olivass V, et al. NF-kappaB activating complex engaged in response to EGFR oncogene inhibition drives tumor cell survival and residual disease in lung cancer. *Cell Rep* 2015; 11:98–110.
153. Imielinski M, Berger AH, Hammerman PS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 2012; 150:1107–20.
154. Steuer CE, Ramalingam SS. Tumor mutation burden: leading immunotherapy to the era of precision medicine? *J Clin Oncol* 2018; 36:631–2.
155. Attarian S, Rahman N, Halmos B. Emerging uses of biomarkers in lung cancer management: molecular mechanisms of resistance. *Ann Transl Med* 2017; 5:377.
156. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2001; 2:127–37.
157. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014; 511:543–50.
158. Marchetti A, Felicioni L, Malatesta S, et al. Clinical features and outcome of patients with non-small-cell lung cancer harboring BRAF mutations. *J Clin Oncol* 2011; 29:3574–9.



159. Planchard D, Besse B, Groen HJM, et al. Dabrafenib plus trametinib in patients with previously treated BRAF(V600E)-mutant metastatic non-small cell lung cancer: an open-label, multicentre phase 2 trial. *Lancet Oncol* 2016;17:984-93.
160. Rangarajan A, Hong SJ, Gifford A, et al. Species- and cell type-specific requirements for cellular transformation. *Cancer Cell* 2004;6:171-83.
161. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014;511:543-50.
162. Kim J, McMillan E, Kim HS, et al. XPO1-dependent nuclear export is a druggable vulnerability in KRAS mutant lung cancer. *Nature* 2016;538:114-7.
163. Wang J, Hu K, Guo J, et al. Suppression of KRAS-mutant cancer through the combined inhibition of KRAS with PLK1 and ROCK. *Nat Commun* 2016;7:11363.
164. Nichols RJ, Haderk F, Stahlhut C, et al. Efficacy of SHP2 phosphatase inhibition in cancers with nucleotide-cycling oncogenic RAS, RAS-GTP dependent oncogenic BRAF and NF1 loss. *BioRxiv*. 2017. Available online: <https://www.biorxiv.org/content/early/2017/09/14/188730>.
165. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014;511:543-50.
166. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 2007;448:561-6.
167. Rotow J, Bivona TG. Understanding and targeting resistance mechanisms in NSCLC. *Nat Rev Cancer* 2017;17:637-58.
168. Peters S, Camidge DR, Shaw AT, et al. Alectinib versus Crizotinib in Untreated ALK-positive non-small-cell lung cancer. *N Engl J Med* 2017;377:829-38.
169. Lin JJ, Kennedy E, Sequist LV, et al. Clinical activity of alectinib in advanced RET-rearranged non-small cell lung cancer. *J Thorac Oncol* 2016;11:2027-32.
170. George J, Lim JS, Jang SJ, et al. Comprehensive genomic profiles of small cell lung cancer. *Nature* 2015;524:47-53.
171. Zhang J, Zhang L, Su X, et al. Translating the therapeutic potential of AZD4547 in FGFR1-amplified non-small cell lung cancer through the use of patient-derived tumor xenograft models. *Clin Cancer Res* 2012;18:6658-67.
172. Paik PK, Shen R, Berger MF, et al. A Phase Ib OpenLabel Multicenter Study of AZD4547 in patients with advanced squamous cell lung cancers. *Clin Cancer Res* 2017;23:5366-73.
173. Nogova L, Sequist LV, Perez Garcia JM, et al. Evaluation of BGJ398, a Fibroblast Growth Factor Receptor 1-3 kinase inhibitor, in patients with advanced solid tumors harboring genetic alterations in fibroblast growth factor receptors: results of a global phase I, dose-escalation and dose-expansion study. *J Clin Oncol* 2017;35:157-65.
174. Lim SH, Sun JM, Choi YL, et al. Efficacy and safety of dovitinib in pretreated patients with advanced squamous non-small cell lung cancer with FGFR1 amplification: A single-arm, phase 2 study. *Cancer* 2016;122:3024-31.
175. Paik PK, Shen R, Berger MF, et al. A Phase Ib OpenLabel Multicenter Study of AZD4547 in patients with advanced squamous cell lung cancers. *Clin Cancer Res* 2017;23:5366-73.
176. Pleasance ED, Cheetham RK, Stephens PJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 2010;463:191-6.
177. Jin G, Kim MJ, Jeon HS, et al. PTEN mutations and relationship to EGFR, ERBB2, KRAS, and TP53 mutations in non-small cell lung cancers. *Lung Cancer* 2010;69:279-83.
178. Paik PK, Shen R, Berger MF, et al. A Phase Ib OpenLabel Multicenter Study of AZD4547 in patients with advanced squamous cell lung cancers. *Clin Cancer Res* 2017;23:5366-73.
179. Xu C, Fillmore CM, Koyama S, et al. Loss of Lkb1 and Pten leads to lung squamous cell carcinoma with elevated PD-L1 expression. *Cancer Cell* 2014;25:590-604.
180. Lastwika KJ, Wilson W 3rd, Li QK, et al. Control of PD-L1 Expression by Oncogenic Activation of the AKT/mTOR pathway in non-small cell lung cancer. *Cancer Res* 2016;76:227-38.
181. Dickson MA, Schwartz GK. Development of cell-cycle inhibitors for cancer therapy. *Curr Oncol* 2009;16:36-43.
182. Nichols RJ, Haderk F, Stahlhut C, et al. Efficacy of SHP2 phosphatase inhibition in cancers with nucleotide-cycling oncogenic RAS, RAS-GTP dependent oncogenic BRAF and NF1 loss. *BioRxiv*. 2017.



183. Farago AF, Le LP, Zheng Z, et al. Durable Clinical Response to Entrectinib in NTRK1-rearranged non-small cell lung cancer. *J Thorac Oncol* 2015;10:1670-4.
184. Kobayashi Y, Mitsudomi T. Not all epidermal growth factor receptor mutations in lung cancer are created equal: Perspectives for individualized treatment strategy. *Cancer Sci* 2016;107:1179-86.
185. Huang Q, Schneeberger VE, Luetke N, et al. Preclinical Modeling of KIF5B-RET Fusion Lung Adenocarcinoma. *Mol Cancer Ther* 2016;15:2521-9.
186. Ramazzotti, D.; Lal, A.; Wang, B.; Batzoglou, S.; Sidow, A. Multi-omic tumor data reveal diversity of molecular mechanisms that correlate with survival. *Nat. Commun.* 2018, 9, 4453. [CrossRef]
187. Chaudhary, K.; Poirion, O.B.; Lu, L.; Garmire, L.X. Deep Learning-Based Multi-Omics Integration Robustly Predicts Survival in Liver Cancer. *Clin. Cancer Res.* 2018, 24, 1248–1259.
188. Yuan, C.; Yang, H. Research on K-Value Selection Method of K-Means Clustering Algorithm. *J—Multidiscip. Sci. J.* 2019, 2, 16.
189. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 1987, 20, 13.
190. Calinski, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat. - Theory Methods* 1974, 3, 27.
191. Laurens van der Maaten, G.H. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* 2008, 9, 2579–2605.
192. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015, 43, e47.
193. Anaya, J. OncoLnc: Linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *Peer J. Comput. Sci.* 2016, 2, 11.
194. Sharma SV, Bell DW, Settleman J, Haber DA (2007) Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 7:169–181.
195. Kosaka T, et al. (2004) Mutations of the epidermal growth factor receptor gene in lung cancer: Biological and clinical implications. *Cancer Res* 64:8919–8923.
196. Yun CH, et al. (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: Mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* 11:217–227.
197. Jura N, et al. (2011) Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Mol Cell* 42:9–22.
198. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptors. *Cell* 125:1137– 1149.
199. Jänne PA, Johnson BE (2006) Effect of epidermal growth factor receptor tyrosine kinase domain mutations on the outcome of patients with non-small cell lung cancer treated with epidermal growth factor receptor tyrosine kinase inhibitors. *Clin Cancer Res* 12:4416s–4420s.
200. Mitsudomi T, Yatabe Y (2007) Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Sci* 98:1817–1824.
201. Sequist LV, et al. (2013) Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *J Clin Oncol* 31: 3327–3334.
202. Cross DA, et al. (2014) AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov* 4:1046–1061.
203. Mok TS, et al. (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947–957.
204. Ramalingam SS, et al. (2018) Osimertinib as first-line treatment of EGFR mutation positive advanced non-small-cell lung cancer. *J Clin Oncol* 36:841–849.
205. Yasuda H, Kobayashi S, Costa DB (2012) EGFR exon 20 insertion mutations in non small-cell lung cancer: Preclinical data and clinical implications. *Lancet Oncol* 13:e23– e31.
206. Kosaka T, et al. (2017) Response heterogeneity of EGFR and HER2 exon 20 insertions to covalent EGFR and HER2 inhibitors. *Cancer Res* 77:2712–2721.
207. Arcila ME, et al. (2013) EGFR exon 20 insertion mutations in lung adenocarcinomas: Prevalence, molecular heterogeneity, and clinicopathologic characteristics. *Mol Cancer Ther* 12:220-229.



208. Ruan Z, Kannan N (2018) Altered conformational landscape and dimerization dependency underpins the activation of EGFR by α C- β 4 loop insertion mutations. *Proc Natl Acad Sci USA* 115:E8162–E8171.
209. Shan Y, Arkhipov A, Kim ET, Pan AC, Shaw DE (2013) Transitions to catalytically inactive conformations in EGFR kinase. *Proc Natl Acad Sci USA* 110:7270–7275.
210. Ruan Z, Katiyar S, Kannan N (2017) Computational and experimental characterization of patient derived mutations reveal an unusual mode of regulatory spine assembly and drug sensitivity in EGFR kinase. *Biochemistry* 56:22–32.
211. Wang DD, Zhou W, Yan H, Wong M, Lee V (2013) Personalized prediction of EGFR mutation-induced drug resistance in lung cancer. *Sci Rep* 3:2855.
212. Ma L, et al. (2015) EGFR mutant structural database: Computationally predicted 3D structures and the corresponding binding free energies with gefitinib and erlotinib. *BMC Bioinformatics* 16:85.
213. Ruiz-Carmona S, et al. (2014) rDock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol* 10:e1003571.
214. Araki M, et al. (2016) The effect of conformational flexibility on binding free energy estimation between kinases and their inhibitors. *J Chem Inf Model* 56:2445–2456.
215. Brown JB, Nakatsui M, Okuno Y (2014) Constructing a foundational platform driven by Japan's K supercomputer for next-generation drug design. *Mol Inform* 33:732–741.
216. Nakaoku T, et al. (2018) A secondary RET mutation in the activation loop conferring resistance to vandetanib. *Nat Commun* 9:625.
217. Katayama R, et al. (2014) Two novel ALK mutations mediate acquired resistance to the next-generation ALK inhibitor alectinib. *Clin Cancer Res* 20:5686–5696.
218. Biankin AV, Piantadosi S, Hollingsworth SJ (2015) Patient-centric trials for therapeutic development in precision oncology. *Nature* 526:361–370.
219. Hirano T, et al. (2015) In vitro modeling to determine mutation specificity of EGFR tyrosine kinase inhibitors against clinically relevant EGFR mutants in non-small-cell lung cancer. *Oncotarget* 6:38789–38803.
220. Wang DD, Zhou W, Yan H, Wong M, Lee V (2013) Personalized prediction of EGFR mutation-induced drug resistance in lung cancer. *Sci Rep* 3:2855.
221. Ma L, et al. (2015) EGFR mutant structural database: Computationally predicted 3D structures and the corresponding binding free energies with gefitinib and erlotinib. *BMC Bioinformatics* 16:85.
222. Ruiz-Carmona S, et al. (2014) rDock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol* 10:e1003571.
223. Ruiz-Carmona S, et al. (2014) rDock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol* 10:e1003571.
224. Miller BR, 3rd, et al. (2012) MMPBSA.py: An efficient program for end-state free energy calculations. *J Chem Theory Comput* 8:3314–3321.
225. Wang DD, Zhou W, Yan H, Wong M, Lee V (2013) Personalized prediction of EGFR mutation-induced drug resistance in lung cancer. *Sci Rep* 3:2855.
226. Ma L, et al. (2015) EGFR mutant structural database: Computationally predicted 3D structures and the corresponding binding free energies with gefitinib and erlotinib. *BMC Bioinformatics* 16:85.

Cite this Article: Purbasha Pati, Pooja Kumari, Neha Kumari, Kumari Surekha Mahto, Shekhar Marandi, Sadaf Naaz, Sweety Guria Rani, Naiyar Naaz, Rupa Verma, Saurabh Kumar Gupta, Mukesh Nitin (2022). Current Advances in Computational biology of Lung Cancer. International Journal of Current Science Research and Review, 5(1), 78-99