

# Identifying Prognostic Groups Using Machine Learning Tools in Patients Undergoing Chemoradiation for Inoperable Locally Advanced Nonsmall Cell Lung Carcinoma

Anjali K. Pahuja<sup>1</sup> Kundan Singh Chufal<sup>1</sup> Irfan Ahmad<sup>1</sup> Ram Bajpai<sup>2</sup> Rajpal Singh<sup>1</sup>  
Rahul Lal Chowdhary<sup>1</sup> Maithili Sharma<sup>1</sup>

<sup>1</sup>Department of Radiation Oncology, Rajiv Gandhi Cancer Institute and Research Centre, New Delhi, India

<sup>2</sup>Primary Care Centre Versus Arthritis, Research Institute for Primary Care and Health Sciences, Keele University, Staffordshire, United Kingdom

**Address for correspondence** Anjali K. Pahuja, MD, Department of Radiation Oncology, Rajiv Gandhi Cancer Institute and Research Centre, Sector-5, Rohini, New Delhi 110085, India (e-mail: anjali\_kakria@yahoo.com).

Asian J Oncol 2019;5:56–63

## Abstract

**Introduction** Unresectable stage III nonsmall cell lung cancer (NSCLC) continues to have dismal 5-year overall survival (OS) rate. However, a subset of the patients treated with chemoradiation show significantly better outcome. Prediction of treatment outcome can be improved by utilizing machine learning tools, such as cluster analysis (CA), and is capable of identifying complex interactions among many variables. We have utilized CA to identify a cluster with good prognosis within stage III NSCLC.

**Materials and Methods** Retrospective analysis of treatment outcomes was done for 92 patients who underwent chemoradiation for inoperable locally advanced NSCLC from 2012 to 2018. Using various patient- and treatment-related variables, an exploratory factor analysis was performed to extract factors with eigenvalue > 1. An appropriate number of homogeneous groups were identified using agglomerative hierarchical cluster analysis. Further K-mean cluster analysis was applied to classify each patient into their homogeneous clusters. The newly formed cluster variable was used as an independent variable to estimate survival over time using Kaplan–Meier method.

**Results** With a median follow-up of 18 months, median OS was 14 months. Using CA, three prognostic clusters were obtained. Cluster 2 with 36 patients had a median OS of 36 months, whereas Cluster 1 with 34 patients had a median OS of 20 months ( $p = 0.004$ ).

**Conclusion** A cluster could thus be identified with a relatively good prognosis within stage III NSCLC. Using CA, we have attempted to create a model which may provide more specific prognostic information in addition to that provided by tumor node metastasis-based models.

## Keywords

- ▶ chemoradiation
- ▶ machine learning
- ▶ nonsmall cell lung carcinoma

## Introduction

The most frequent cause of cancer-related mortality worldwide remains to be lung cancer despite progress made in all oncological modalities involved in its management. Nearly 85% of all lung neoplasia are classified as nonsmall cell lung

cancer (NSCLC)<sup>1</sup> of whom about one-third are diagnosed in a locally advanced stage.<sup>2</sup> Concomitant chemoradiotherapy (CCRT) is the established standard of care for unresectable stage III disease with evidence of better results compared with either modality used alone or sequentially.<sup>3</sup> Despite the

received  
October 24, 2019  
accepted  
November 5, 2019

DOI <https://doi.org/10.1055/s-0039-3401437>  
ISSN 2454-6798.

©2019 Spring Hope Cancer Foundation & Young Oncologist Group of Asia

License terms



dismal 5-year overall survival (OS) rate of 15 to 35% for stage IIIA and 5 to 10% for stage IIIB,<sup>4</sup> a subset of patients in these stages show significantly better OS. Conventional patient, tumor, and treatment related parameters often do not correlate with the survival outcome, largely due to complex interactions between tumor biology, tumor microenvironment, radiation dosimetry, and patient-related variables.

Prediction of treatment outcome can be improved by utilizing machine learning (ML) tools capable of identifying complex interactions among variables. Of the many such tools available, cluster analysis (CA) provides a potential relationship and constructs a systematic structure in large number of variables and observations. K-means clustering is an unsupervised learning algorithm that tries to cluster data based on its similarity. We have utilized K-mean CA to identify a cluster with good prognosis within unresectable stage III NSCLC with significantly better overall survival.

## Materials and Methods

Retrospective analysis of treatment outcomes was done for patients who underwent definite chemoradiation upfront or sequentially after neoadjuvant chemotherapy (NACT) for inoperable locally advanced NSCLC, adenocarcinoma (AC), and squamous cell carcinoma (SCC) from 2012 to 2018.

Patients with small cell or histology other than AC or SCC, history of previous thoracic radiotherapy and presence of a second primary were excluded from the study. All the patients were restaged using the latest eighth edition of American Joint Commission on Cancer (AJCC).<sup>5</sup>

Patient- and treatment-related data were retrieved from hospital medical records, available in the form of hospital case files, electronic medical records, and radiotherapy cards. Patients or their caregivers were contacted telephonically for obtaining the latest survival status.

## Statistical Analysis

Age has been summarized via mean and standard deviation (SD), and radiotherapy dose and overall treatment time by median and interquartile range (IQR), and categorical variables via frequencies and percentages (► **Table 1**).

We first explored any underlying factor structure between the study variables (age, gender, comorbidity, positron emission tomography-complete response, radiotherapy dose, overall treatment time (OTT), radiotherapy technique, and staging). For this purpose, exploratory factor analysis was performed to extract factors using oblique rotation (direct oblimin, which assumes correlated factors), with eigenvalue > 1 criterion. A particular variable loaded a factor if the value of the loading was highest for that factor. Between variables' correlation matrix was initially visually inspected for extreme multicollinearity. Subsequently the Kaiser-Meyer-Olkin test was performed to assess the adequacy of data for factor analysis (test values of >0.5 as acceptable), and significance of Bartlett's test (testing the null hypothesis that the correlation matrix is an identity matrix).<sup>6</sup> Sensitivity analysis was also done to check any variations by considering orthogonal varimax rotation which does not

assume correlated factors. Next, agglomerative hierarchical cluster analysis was performed using Ward's minimum variance method to identify the appropriate number of homogeneous groups of patients based on the same set as in the factor analysis.<sup>7</sup> Then, these identified numbers were used in the K-mean cluster analysis to classify each patient into their homogeneous clusters. The newly formed cluster variable is used as independent variable to estimate survival over time using Kaplan-Meier method and Cox's model was used to estimate relative hazards (hazard ratio [HR] with 95% confidence interval [CI]). As a complementary analysis, a Cox's model was also constructed using full set of predictor variables to understand their independent effect on survival.

## Results

Of the 152 patients presenting to the department of radiation oncology for definite radiotherapy, data of 92 patients who completed their prescribed treatment and presented for the first follow-up was analyzed for descriptive statistics (► **Fig. 1**).

Of the 92 patients who received definite chemoradiation for unresectable locoregionally advanced NSCLC, 82.6% were males. Among the patients, 45.7% had AC while the remaining had SCC. The mean dose of radiation was 60.84 Gy (50–70 Gy). Conformal techniques, such as intensity modulated radiotherapy (IMRT) or image guided radiotherapy (IGRT), were employed in the treatment of 72.8% of the patients. The mean OTT in the 92 patients analyzed for survival outcomes was 90.9 days (range: 40–373 days). Other treatment related characteristics have been shown in ► **Table 1**.

Median follow-up for all patients was 18 months (SD ± 11.5; range 3–59 months). At last follow-up, among 92 patients, 27 (30.5%) patients were alive without disease, 22 (26.8%) were alive with disease, and 37 (42%) patients had died with disease. Median recurrence free survival (RFS) and OS were 14 months and 22 months, respectively and 2-year RFS and OS were 14.4 and 53.3%, respectively, across all the stages (► **Fig. 2**).

Cluster 2 comprising of 36 (39.1%) patients treated with IMRT/IGRT, up to a mean dose of 64 Gy (range: 62.3–65.9 Gy) and with a mean OTT of 49 days (range: 45.3–90.8 days), had a median OS of 36 months, whereas, cluster 3 comprising of 22 (23.9%) patients treated with conventional radiation techniques (two-dimensional [2D]), up to a mean dose of 58 Gy (range: 53.5–63.7 Gy) with a mean OTT of 76 days (range: 44–146 days), had a median OS of 19 months ( $p = 0.000$ ). We could thus identify a cluster with good prognosis within stage III where an adequate radiation dose (60–66 Gy) using an improved radiation technique delivered in shorter OTT (<90 days) was associated with better overall survival (► **Fig. 3**).

## Discussion

The current standard of care in unresectable stage III NSCLC patients is curative-dose radiotherapy along with platinum-based chemotherapy. However, only 15 to 30% of patients survive at 5 years, corresponding to a median

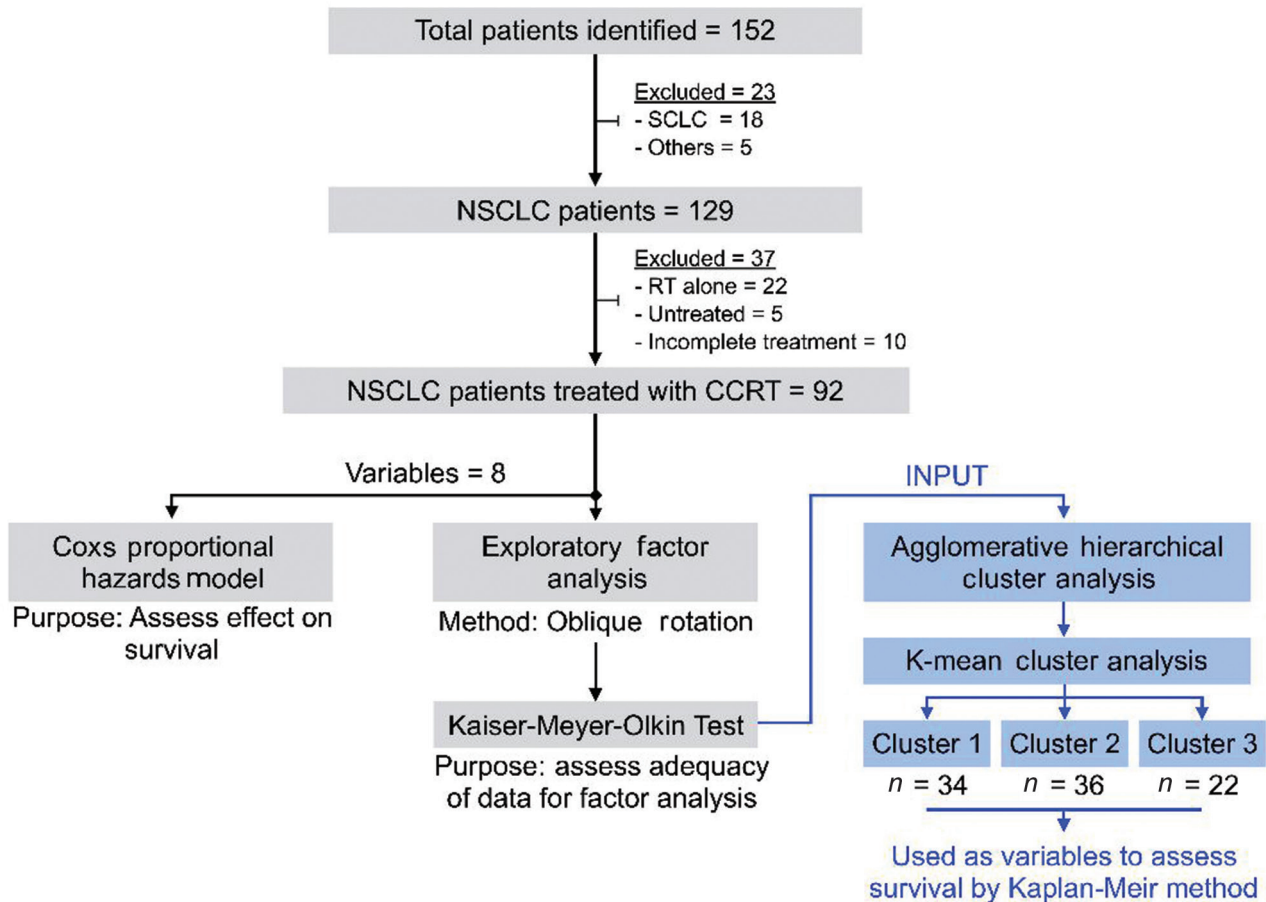
**Table 1** Patient and treatment characteristics

| Parameters                   | n  | Mean or percentage         | SD    | Range  |
|------------------------------|----|----------------------------|-------|--------|
| Age (y)                      | 92 | 60.73 (95% CI: 58.7–62.6)  | 9.343 | 39–83  |
| <b>Gender</b>                |    |                            |       |        |
| Male                         | 76 | 82.6                       |       |        |
| Female                       | 16 | 17.4                       |       |        |
| <b>Comorbidity</b>           |    |                            |       |        |
| Yes                          | 43 | 46.7                       |       |        |
| No                           | 49 | 53.3                       |       |        |
| <b>Histology</b>             |    |                            |       |        |
| ACC                          | 42 | 45.7                       |       |        |
| SCC                          | 50 | 54.3                       |       |        |
| <b>Stage</b>                 |    |                            |       |        |
| IIIA                         | 13 | 14.1                       |       |        |
| IIIB                         | 22 | 23.9                       |       |        |
| IIIC                         | 57 | 62.0                       |       |        |
| Radiation dose               | 92 | 60.84 (95% CI: 59.36–61.6) | 4.89  | 50–70  |
| <b>CT timing</b>             |    |                            |       |        |
| NACT                         | 47 | 51.1                       |       |        |
| CCRT                         | 45 | 48.9                       |       |        |
| NACT cycles (n = 47)         |    |                            |       |        |
| ≤3cycles                     | 29 | 61.7                       |       |        |
| >3cycles                     | 18 | 38.3                       |       |        |
| <b>CT regimen for NACT</b>   |    |                            |       |        |
| Platinum + taxane            | 23 | 32.9                       |       |        |
| Platinum + etoposide         | 23 | 32.9                       |       |        |
| Platinum + gemcitabine       | 7  | 10.0                       |       |        |
| Platinum + pemetrexate       | 11 | 15.7                       |       |        |
| Others                       | 6  | 8.5                        |       |        |
| <b>CCRT (n = 45)</b>         |    |                            |       |        |
| ≥4 cycles                    | 18 | 40                         |       |        |
| <4 cycles                    | 27 | 60                         |       |        |
| TTT (d)                      | 92 | 46.9 (95% CI: 45.9–48.0)   | 5.1   | 38–66  |
| OTT (d)                      | 92 | 90.9 (95% CI: 79.2–102.8)  | 56.9  | 40–373 |
| <b>RT technique (n = 92)</b> |    |                            |       |        |
| 2D technique                 | 25 | 27.2                       |       |        |
| IMRT/IGRT                    | 67 | 72.8                       |       |        |

Abbreviations: 2D, two-dimensional; ACC, adenocarcinoma; CCRT, concurrent chemoradiotherapy; CT, chemotherapy; IMRT/IGRT, intensity modulated radiotherapy/image guided radiotherapy; NACT, neoadjuvant chemotherapy; OTT, overall treatment time from initiation of any treatment till end of treatment; RT, radiotherapy; SCC, squamous cell carcinoma; SD, standard deviation; TTT, total duration of radiation.

survival of approximately 28 months.<sup>2</sup> Even with the addition of systemic therapy after achieving locoregional control with chemoradiotherapy (CRT), median survival ranges from 18 to 23 months.<sup>8</sup> PACIFIC study that compared consolidation using durvalumab (a human Ig [immunoglobulin] G1 monoclonal antibody that blocks PD-L1 binding to PD-1 and CD80) with

placebo after concurrent chemoradiation in patients with stage III, unresectable NSCLC, demonstrated a 24-month OS rate of 66.3% in favor of durvalumab. At a median follow-up of 25.2 months, the median OS had not been reached at the time of publication of the results.<sup>9</sup> Thus a new standard of care has been established for unresectable NSCLC.



**Fig. 1** Patient selection criteria and selection of variables for cluster analysis. CCRT, concurrent chemoradiotherapy; NSCLC, nonsmall cell lung cancer; RT, radiotherapy.

There has been an attempt at finding the prognostic and predictive factors for locoregionally advanced unresectable NSCLC which may help in proper patient selection for any form of aggressive therapeutic approach to maximize the benefits of treatment in an otherwise dismal scenario. Age, gender, performance status, weight loss in the period of 3 months before diagnosis, baseline hemoglobin value, normal leukocytosis and normal neutrophil count, lactate dehydrogenase (LDH) level, hypercalcemia, hypoalbuminemia, tumor dimension, and involved lymph node burden are some of the known prognostic factors in NSCLC.<sup>10,11</sup> Prognostic classifications have been attempted based on gene signatures in resected specimen such as the 15-gene signatures, independent from stage with an overall HR of 15.02 (95% CI: 5.12–44.04) with consistent results in stages I and II.<sup>12</sup>

However, till date, a correlation between the above mentioned and many other possibly unknown factors has not been established with varied survival outcomes for patients within a given stage. Clearly, better models based on early assessment of response after definite CRT are needed to predict outcome, in time for treatment intensification with additional radiation, early addition of systemic therapy, or application of a different treatment modality.

With an ever-increasing number of patients with locoregionally advanced NSCLC, usually not amenable to surgery,

the data, required to generate prognostic models for risk stratification and treatment intensification has been growing exponentially, making it imperative to utilize the applications of ML. ML has the potential to change the way radiation oncologists follow patients treated with definitive radiotherapy. It is a computerized approach to identify complex mathematical associations within a set of observational data. It is also an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Further, ML may be supervised (prediction, classification, etc.), unsupervised (clustering, probability distribution estimation, etc.), or reinforcement learning (robot, chess machines, etc.). Many questions in oncology can be answered through ML algorithms that aid in decision making in an era of individualized care plans. Within radiation oncology, some work has been done to model individual radiation sensitivity to individualize and adapt therapy.<sup>13</sup>

Further, in the field of radiation for malignancy of lung, early work was confined to predict tumor motion and predicting the need for replanning.<sup>14</sup> Deep neural networks have been used to predict the need for treatment adaptation for lung patients.<sup>15</sup> Evaluation of 32 clinical features per patient in a cohort of 203 stages II and III unresectable NSCLC patients treated with definitive chemoradiation using ML established

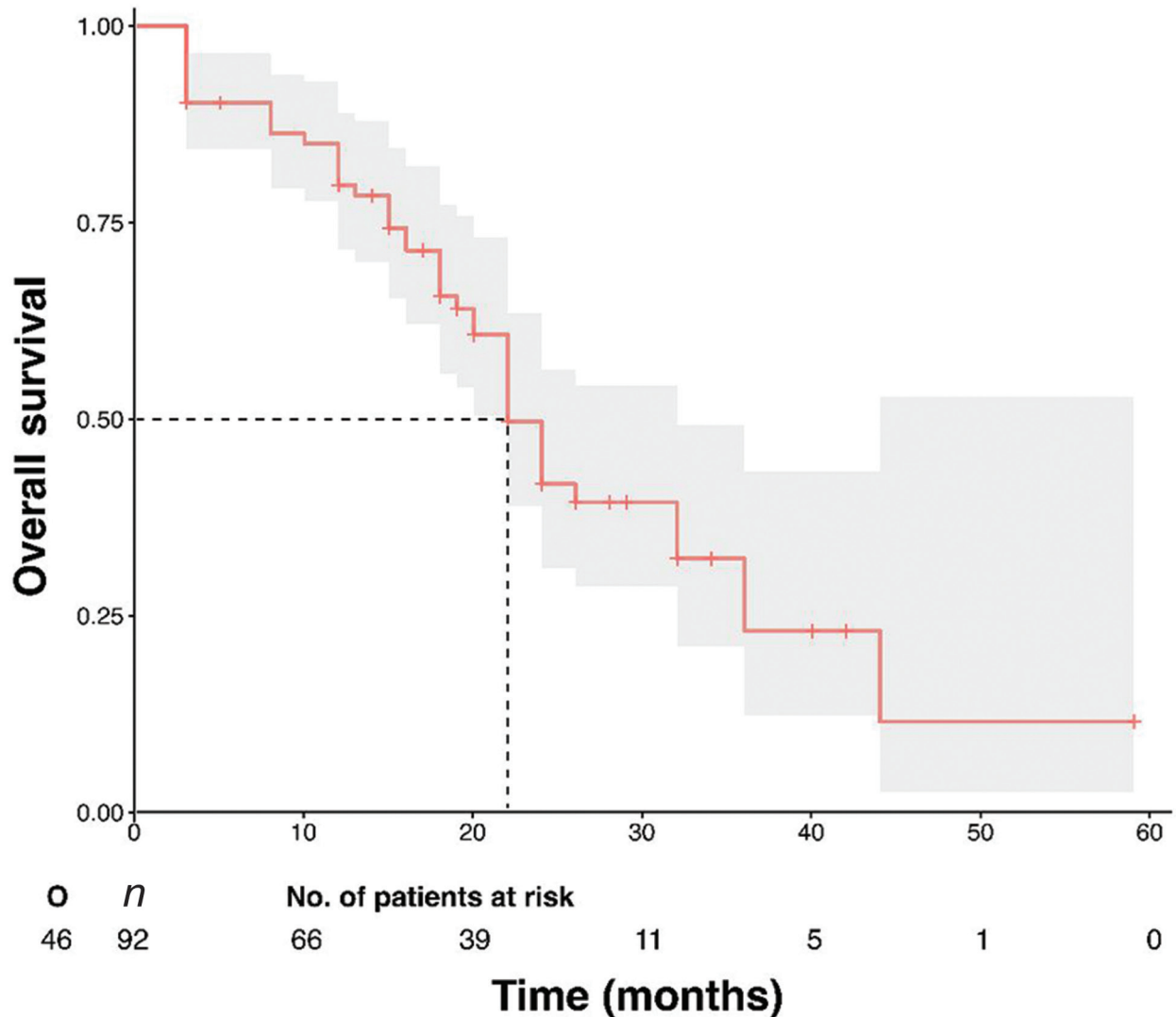


Fig. 2 Kaplan–Meier estimate for OS of 92 patients. OS, overall survival.

random forest as an accurate method to identify known and new predictors of symptomatic radiation pneumonitis.<sup>16</sup>

Analysis of our institute's data of 92 unresectable stage III NSCLC patients who underwent definite chemoradiation revealed a median OS of 26 months and a 2-year OS of 53.3%, consistent with the known survival rates in other series.<sup>2</sup> This cohort of patients belonged to different stage subgroups within stage III (►Table 1), treated with different radiotherapy techniques (at the treating physician's discretion, availability of technology, or patient related logistic constraints), varying tumor doses, over a range of treatment time and variable positron emission tomography (PET) response after CRT. Using multivariable Cox's regression model, the impact of individual variables on survival outcome could be calculated as shown in ►Table 2.

However, the correlation between these independent variables and its impact, if any, on survival required the application of ML tools. We therefore used a clustering method available with SPSS, namely, cluster analysis to derive prognostic groups within stage III. CA is a multivariate method

which aims to classify a sample of patients (or objects) on the basis of a set of measured variables into several different groups such that similar patients are placed in the same group. K-means clustering is a nonhierarchical method of CA in which the desired number of clusters is specified in advance and the "best" solution is chosen. It tends to be used when large datasets are involved.<sup>17</sup>

The ML tool, utilizing eight variables as input, arrived at three prognostic clusters.

Cluster 2 having 36 patients, all of whom had stage IIC NSCLC, were treated with conformal advanced techniques to a dose of >60 Gy and completed their treatment within 90 days. This cluster had a median OS of 36 months which was significantly better than that of patients in cluster 3 ( $p = 0.000$ ) and cluster 1 ( $p = 0.004$ ) with similar stage but different treatment parameters. This was corroborated by the Cox model (►Table 3).

More recently, early tumor shrinkage during the course of concurrent chemoradiation has been proposed as a prognostic factor in stage III.<sup>18</sup> An extensive review of application of ML

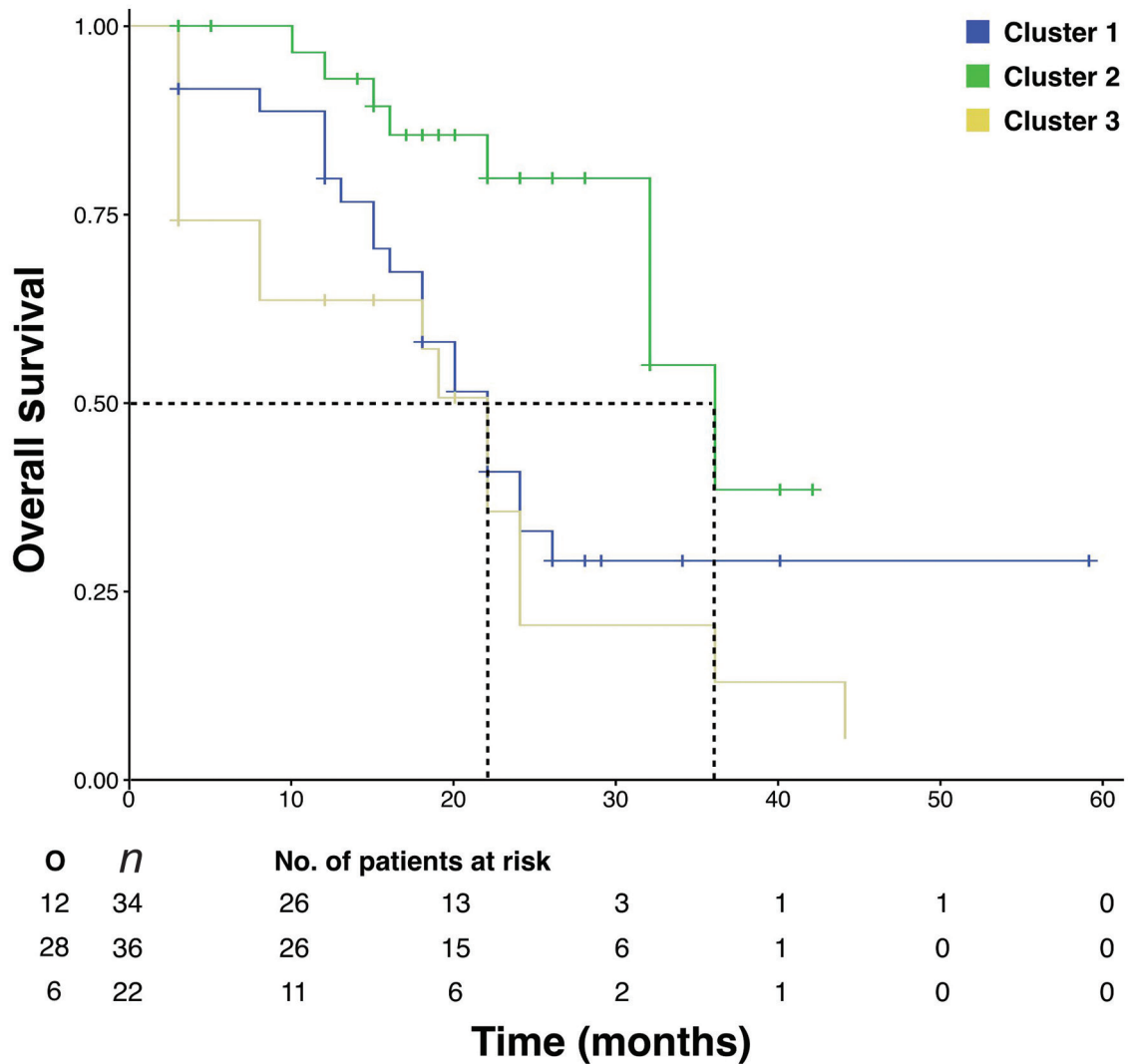


Fig. 3 Kaplan–Meier estimate showing significantly better overall survival for cluster 2 compared with clusters 1 and 3.

in radiotherapy for NSCLC has been published, approaching the radiotherapy process from a workflow perspective, identifying specific areas where a data-centric approach using ML could improve the quality and efficiency of patient care. While touching upon nearly every aspect from patient assessment, simulation, planning, quality assurance, treatment delivery to follow-up, it serves as a guide for clinicians to discuss issues that must be addressed in a timely manner, outside the conventional factors.<sup>19</sup> But there is a paucity of data establishing radiation dose and technique as prognostic factors in a scenario where radiation forms the backbone of treatment and outcome.

RTOG 0617 trial, comparing 60 Gy with 74 Gy, showed a significantly lower survival in the high-dose arm,<sup>20</sup> conflicting with the results of some prospective studies suggesting better local control and higher survival rates with higher radiation doses.<sup>21,22</sup> The unexpected findings of RTOG 0617 have been attributed to prolonged OTT and increased heart toxicity in the high-dose group. A decrease in tumor control probability of 1.6% per day after a 6-week duration of radiation therapy<sup>23</sup> and a 2.0% increase in the risk of death for each day of prolongation in therapy<sup>24</sup> have been published.

Ours being a retrospective analysis, with a wide range of doses delivered (40–70 Gy), over a wide range of OTT, using different techniques of radiation delivery (as per the treating physician’s discretion), it would have been difficult to reach a definite conclusion with conventional statistical methods. However, with the help of CA, the same heterogeneity could be utilized to conclude that not only the total dose but the technique used to deliver the same and the duration in which it is delivered have an impact on the survival outcome.

A multivariate predictive model, using data from 548 patients with stage III NSCLC, consisting of age, gender, performance status, overall treatment time, equivalent radiation dose, number of positive lymph node stations, and gross-tumor volume has been envisaged as a first building block for a decision support system to predict survival probability for an individual patient with stage III NSCLC.<sup>25</sup> This model was based on patients treated with three-dimensional CRT or IMRT, therefore predictions for patients treated with other techniques were not possible. In the present study, we have been able to incorporate treatment technique that ranged from 2D to IGRT, in the model.



**Table 2** Factors associated with the survival using multivariable Cox's regression model

| Variables                     | Overall survival |           |         | Progression-free survival |            |         |
|-------------------------------|------------------|-----------|---------|---------------------------|------------|---------|
|                               | aHR              | 95% CI    | p-Value | aHR                       | 95% CI     | p-Value |
| Age (y)                       | 1.02             | 0.99–1.05 | 0.210   | 1.00                      | 0.98–1.03  | 0.997   |
| <b>Gender</b>                 |                  |           |         |                           |            |         |
| Male                          | 1.69             | 0.67–4.25 | 0.262   | 0.93                      | 0.48–1.79  | 0.827   |
| Female                        | 1 (Ref.)         |           |         | 1 (Ref.)                  |            |         |
| <b>Any comorbidity</b>        |                  |           |         |                           |            |         |
| Yes                           | 1.61             | 0.79–3.27 | 0.187   | 1.24                      | 0.75–2.06  | 0.398   |
| No                            | 1 (Ref.)         |           |         | 1 (Ref.)                  |            |         |
| <b>PET-CR</b>                 |                  |           |         |                           |            |         |
| Yes                           | 1 (Ref.)         |           |         | 1 (Ref.)                  |            |         |
| No                            | 2.19             | 0.88–5.43 | 0.091   | 5.11                      | 2.01–13.00 | 0.001   |
| <b>Stage</b>                  |                  |           |         |                           |            |         |
| IIIA                          | 1 (Ref.)         |           |         | 1 (Ref.)                  |            |         |
| IIIB                          | 0.22             | 0.05–0.97 | 0.046   | 0.46                      | 0.20–1.06  | 0.069   |
| IIIC                          | 1.26             | 0.48–3.28 | 0.641   | 0.83                      | 0.42–1.62  | 0.579   |
| <b>Radiotherapy technique</b> |                  |           |         |                           |            |         |
| 2D                            | 1 (Ref.)         |           |         | 1 (Ref.)                  |            |         |
| IMRT                          | 0.42             | 0.20–0.86 | 0.018   | 0.55                      | 0.31–0.97  | 0.038   |
| Radiotherapy dose             | 0.87             | 0.80–0.94 | 0.000   | 1.02                      | 0.96–1.08  | 0.488   |
| Log (overall treatment time)  | 1.91             | 1.11–3.32 | 0.020   | 1.38                      | 0.89–2.13  | 0.153   |

Abbreviations: 2D, two-dimensional; aHR, adjusted hazard ratio; CI, confidence interval; IMRT, intensity modulated radiotherapy; PET-CR, positron emission tomography-complete response; Ref. reference.

**Table 3** Clusters associated with the survival using univariable Cox's regression model

| Clusters  | Progression-free survival |           |         | Overall survival |            |         |
|-----------|---------------------------|-----------|---------|------------------|------------|---------|
|           | HR                        | 95% CI    | p-Value | HR               | 95% CI     | p-Value |
| Cluster 1 | 0.82                      | 0.48–1.41 | 0.466   | 2.98             | 1.32–6.71  | 0.008   |
| Cluster 2 | 1 (Ref.)                  |           |         | 1 (Ref.)         |            |         |
| Cluster 3 | 1.51                      | 0.85–2.68 | 0.159   | 4.27             | 1.82–10.03 | 0.001   |

Abbreviations: HR, hazard ratio; CI, confidence interval; Ref. reference.

### Limitations

In addition to the retrospective nature of the study, other limitations include small sample size and lack of external validation. However, this is a real-world data, highlighting treatment heterogeneity, in the absence of any randomization. Using cluster analysis, we have attempted to create a model which may provide clinicians with more specific prognostic information in addition to that provided by tumor node metastasis (TNM)-based models.

### Conclusion

Apparently, a homogenous group according to the TNM staging system, patients with stage III NSCLC forms a heterogeneous group, as reflected in the survival outcome. Instead of focusing on a handful of variables in individual studies, large databases should be integrated to design prediction models.

While still in its stage of infancy, we envisage that data sharing together with machine learning tools can provide something much better than conventional statistical methods in the near future.

### Conflict of Interest

None declared.

### References

- 1 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66(1):7–30
- 2 Yoon SM, Shaikh T, Hallman M. Therapeutic management options for stage III non-small cell lung cancer. *World J Clin Oncol* 2017;8(1):1–20
- 3 Le Chevalier T, Arriagada R, Quoix E, et al. Radiotherapy alone versus combined chemotherapy and radiotherapy in unresectable non-small cell lung carcinoma. *Lung Cancer* 1994;10(Suppl 1):S239–S244

- 4 Antoni D, Mornex F. Chemoradiotherapy of locally advanced nonsmall cell lung cancer: state of the art and perspectives. *Curr Opin Oncol* 2016;28(2):104–109
- 5 Amin MB, Edge S, Greene F, et al, Eds. *AJCC Cancer Staging Manual* (8th edition). Springer International Publishing: American Joint Commission on Cancer 2017
- 6 Ferguson E, Cox T. Exploratory factor analysis: a users' guide. *Int J Sel Assess* 1993;1:84–94
- 7 Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 1963;58:236–244
- 8 Skrzypski M, Jassem J. Consolidation systemic treatment after radiochemotherapy for unresectable stage III non-small cell lung cancer. *Cancer Treat Rev* 2018;66:114–121
- 9 Antonia SJ, Villegas A, Daniel D, et al; PACIFIC Investigators. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *N Engl J Med* 2017;377(20):1919–1929
- 10 Crvenkova S. Prognostic factors and survival in non-small cell lung cancer patients treated with chemoradiotherapy. *Open Access Maced J Med Sci* 2015;3(1):75–79
- 11 Marianne Paesmans. Prognostic and predictive factors for lung cancer. *Breathe (Sheff)* 2012;9:112–121
- 12 Zhu CQ, Ding K, Strumpf D, et al. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J Clin Oncol* 2010;28(29):4417–4424
- 13 Feng M, Suresh K, Schipper MJ, et al. Individualized adaptive stereotactic body radiotherapy for liver tumors in patients at high risk for liver damage: a phase 2 clinical trial. *JAMA Oncol* 2018;4(1):40–47
- 14 Ruan D, Keall P. Online prediction of respiratory motion: multidimensional processing with low-dimensional feature learning. *Phys Med Biol* 2010;55(11):3011–3025
- 15 Tseng HH, Luo Y, Cui S, Chien JT, Ten Haken RK, Naqa IE. Deep reinforcement learning for automated radiation adaptation in lung cancer. *Med Phys* 2017;44(12):6690–6705
- 16 Luna JM, Chao HH, Diffenderfer ES, et al. Predicting radiation pneumonitis in locally advanced stage II-III non-small cell lung cancer using machine learning. *Radiother Oncol* 2019;133:106–112
- 17 Cornish R. Statistics: cluster analysis. Available at: <http://www.statstutor.ac.uk/resources/uploaded/clusteranalysis.pdf> Accessed on May 24, 2019
- 18 Wei M, Ye Q, Wang X, et al. Early tumor shrinkage served as a prognostic factor for patients with stage III non-small cell lung cancer treated with concurrent chemoradiotherapy. *Medicine (Baltimore)* 2018;97(19):e0632
- 19 Feng M, Valdes G, Dixit N, Solberg TD. Machine learning in radiation oncology: opportunities, requirements, and needs. *Front Oncol* 2018;8:110
- 20 Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): a randomised, two-by-two factorial phase 3 study. *Lancet Oncol* 2015;16(2):187–199
- 21 van Baardwijk A, Reymen B, Wanders S, et al. Mature results of a phase II trial on individualised accelerated radiotherapy based on normal tissue constraints in concurrent chemoradiation for stage III non-small cell lung cancer. *Eur J Cancer* 2012;48(15):2339–2346
- 22 Rengan R, Rosenzweig KE, Venkatraman E, et al. Improved local control with higher doses of radiation in large-volume stage III non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2004;60(3):741–747
- 23 Cox JD, Pajak TF, Asbell S, et al. Interruptions of high-dose radiation therapy decrease long-term survival of favorable patients with unresectable non-small cell carcinoma of the lung: analysis of 1244 cases from 3 Radiation Therapy Oncology Group (RTOG) trials. *Int J Radiat Oncol Biol Phys* 1993;27(3):493–498
- 24 Machtay M, Hsu C, Komaki R, et al. Effect of overall treatment time on outcomes after concurrent chemoradiation for locally advanced non-small-cell lung carcinoma: analysis of the Radiation Therapy Oncology Group (RTOG) experience. *Int J Radiat Oncol Biol Phys* 2005;63(3):667–671
- 25 Oberije C, De Ruyscher D, Houben R, et al. A validated prediction model for overall survival from stage III non-small cell lung cancer: toward survival prediction for individual patients. *Int J Radiat Oncol Biol Phys* 2015;92(4):935–944