PSI workshop conference

Reweighting randomized controlled trial (RCT) evidence to better reflect real life – a case study of the Innovation in Medicine initiative using patients with non-small cell lung cancer (NSCLC)

Alan J M Brnabic
Principal Research Scientist
Eli Lilly
September 2018
Disclaimer

The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of Eli Lilly.
Acknowledgements

Abrams KR\textsuperscript{2}  
Belger M\textsuperscript{1}  
Didden EM\textsuperscript{5}  
Efthimiou O\textsuperscript{4}  
Faries D\textsuperscript{1}  
Girvan A\textsuperscript{1}  
Happich M\textsuperscript{1}  
Jonsson P\textsuperscript{3}  
Johnston J\textsuperscript{1}  
Kadziola Z\textsuperscript{1}  
Ruffieux Y\textsuperscript{4}  
Winfree K\textsuperscript{1}

\textsuperscript{1}Eli Lilly, Indianapolis, IN, USA,  
\textsuperscript{2}University of Leicester, Leicester, UK,  
\textsuperscript{3}National Institute for Health and Care Excellence (NICE), Manchester, UK  
\textsuperscript{4}Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland  
\textsuperscript{5}F.Hoffmann-La Roche Ltd, Basel, Switzerland
Abstract

The objective of the presentation will be to present a case study that assesses the generalizability of efficacy (overall survival [OS]) from the pivotal RCT (JMDB) comparing pemetrexed with gemcitabine to treat non-squamous non-small cell lung cancer using real-world data from a prospective observational study (FRAME) using a reweighting approach. Both inverse propensity scoring and entropy balancing were used to reweight the RCT data based on the real-world FRAME data in an attempt to mirror routine clinical practice in the trial setting.
Presentation contents

♦ Background to studies
♦ Summary of approach
♦ Findings
♦ Recommendations from IMI
♦ Further research/ challenges
Background

- Although the demonstration of improved patient and clinical outcomes within randomized controlled trials (RCTs) is widely accepted as foundational evidence of the efficacy of new treatments, concerns are frequently expressed that RCTs lack external validity.

- Health technology assessment (HTA) are frequently questioning the generalisability of results from RCTS to routine clinical practice.

- GetReal, a project under the umbrella of the Innovation in Medicine Initiative, has explored how “real-life” clinical data can be brought in earlier in drug development.

- We describe a case study that considers lung cancer, the most common cancer worldwide and investigate the generalizability of overall survival.
Case Study: Studies Included – RCT - JMDB

Case Study: Studies Included - RWE- observational study

**FRAME**

- Only baseline data is needed from the RWD source

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Pem + Plat</th>
<th>Gem + Plat</th>
<th>Tax + Plat</th>
<th>Vin + Plat</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any histology NSCLC and any platinum&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>n = 569</td>
<td>n = 360</td>
<td>n = 295</td>
<td>n = 300</td>
<td>n = 1564</td>
</tr>
<tr>
<td>Median age (range) in years ≥70 yrs, %</td>
<td>62 (33-86)</td>
<td>65 (38-84)</td>
<td>65 (37-87)</td>
<td>64 (34-83)</td>
<td>64 (33-87)</td>
</tr>
<tr>
<td>Gender male, %</td>
<td>23.4</td>
<td>34.7</td>
<td>36.3</td>
<td>28.0</td>
<td>29.3</td>
</tr>
<tr>
<td>Never smoked&lt;sup&gt;c&lt;/sup&gt;, %</td>
<td>66.6</td>
<td>78.3</td>
<td>73.9</td>
<td>71.0</td>
<td>71.7</td>
</tr>
<tr>
<td>ECOG PS 2-3, %</td>
<td>13.7</td>
<td>7.8</td>
<td>5.9</td>
<td>8.3</td>
<td>10.4</td>
</tr>
<tr>
<td>Disease stage IV, %</td>
<td>17.9</td>
<td>11.1</td>
<td>22.7</td>
<td>13.0</td>
<td>16.8</td>
</tr>
<tr>
<td>Non-squamous NSCLC and any platinum&lt;sup&gt;b&lt;/sup&gt;</td>
<td>n = 553</td>
<td>n = 201</td>
<td>n = 189</td>
<td>n = 160</td>
<td>n = 1130</td>
</tr>
<tr>
<td>Median age (range) in years ≥70 yrs, %</td>
<td>62 (33-86)</td>
<td>66 (38-84)</td>
<td>64 (37-85)</td>
<td>63 (34-81)</td>
<td>63 (31-86)</td>
</tr>
<tr>
<td>Gender male, %</td>
<td>23.0</td>
<td>31.8</td>
<td>33.3</td>
<td>25.0</td>
<td>26.5</td>
</tr>
<tr>
<td>Never smoked&lt;sup&gt;c&lt;/sup&gt;, %</td>
<td>66.4</td>
<td>73.1</td>
<td>69.8</td>
<td>64.4</td>
<td>68.0</td>
</tr>
<tr>
<td>ECOG PS 2-3, %</td>
<td>14.1</td>
<td>10.9</td>
<td>12.7</td>
<td>8.8</td>
<td>12.5</td>
</tr>
<tr>
<td>Disease stage IV, %</td>
<td>18.1</td>
<td>9.5</td>
<td>21.7</td>
<td>17.5</td>
<td>17.3</td>
</tr>
<tr>
<td>Non-squamous NSCLC and cisplatin</td>
<td>n = 374</td>
<td>n = 107</td>
<td>n = 44</td>
<td>n = 91</td>
<td>n = 633&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Median age (range) in years ≥70 yrs, %</td>
<td>61 (33-86)</td>
<td>63 (38-79)</td>
<td>60 (37-76)</td>
<td>59 (38-81)</td>
<td>60 (33-86)</td>
</tr>
<tr>
<td>Gender male, %</td>
<td>12.8</td>
<td>14.0</td>
<td>9.1</td>
<td>15.4</td>
<td>13.0</td>
</tr>
<tr>
<td>Never smoked&lt;sup&gt;c&lt;/sup&gt;, %</td>
<td>66.3</td>
<td>72.9</td>
<td>63.6</td>
<td>62.6</td>
<td>66.8</td>
</tr>
<tr>
<td>ECOG PS 2-3, %</td>
<td>13.9</td>
<td>11.2</td>
<td>9.1</td>
<td>13.2</td>
<td>12.8</td>
</tr>
<tr>
<td>Disease stage IV, %</td>
<td>13.6</td>
<td>5.6</td>
<td>22.7</td>
<td>13.2</td>
<td>13.1</td>
</tr>
</tbody>
</table>
Case Study: Methods

♦ 2-stage approach

1. the baseline profile of patients in both studies was compared to determine common variables to be used in the weighting

2. Weight the RCT outcome of interest (overall survival) based on the list of variables identified in the first step

♦ Patients were weighted in JMDB using two methods:

1. the inverse propensity score (IPS)\(^1\)

2. entropy balancing (EB) method.\(^2\)

---


\(^2\)Hainmueller Political Analysis 2012 20(1); 25-46
Reweights the RCT patients so as to match the RW population

For example if the cohort patients were found to be significantly older than those participating in the RCT, then the method would tend to downweight the younger patients in the RCT.

Weight RCT outcome using an *algorithm* (IPS or Entropy Balancing) to *match* with RWD data on selected baseline characteristics
Case Study: IPS

♦ The IPS method utilizes a logistic model with study membership as the outcome and the covariates identified as independent variables. \( Y_i = 1 \) if patient \( i \) was an RCT patient, \( Y_i = 0 \) otherwise.

♦ Using the results of this logistic regression, we estimated a propensity score \( \hat{p}_i \) for each patient \( i \).

♦ This propensity score corresponds to the probability of this patient having participated in the RCT rather than in the observational studies, given the set of balancing covariates.

♦ We then assigned the following weight to patient \( i \) in the RCT:

\[
w_i = \frac{1 - \hat{p}_i}{\hat{p}_i}
\]

♦ We kept in our analyses only the reweighted RCT patients.

---


2. Hainmueller Political Analysis 2012 20(1); 25-46
Case Study: Methods

- The propensity adjusted balance between the studies (JMDB vs FRAME) and between the weighted treatments cohorts within JMDB were examined using standardized differences.

- Total patient characteristics in the FRAME study were used to reweight the total non-squamous population in the JMDB trial. (We applied the reweighing independently of the treatment, thereby preserving treatment randomization).

- Overall survival was estimated using Cox proportional hazard models and implemented via the weight statement in both PROC PHREG and PROC LIFETEST.

- Because of the uncertainty around the calculated standard errors, 1000 Bootstrap samples with replacement of each of the datasets (using sample size from original datasets) were created.

- Weights described above were calculated for each dataset and applied to each patient in JMDB for the analysis using the weight statement.

- A distribution of hazard ratios and associated 95% confidence intervals (CIs) was obtained using the percentile approach (ie, selecting the 2.5 and 97.5 percentiles of the bootstrap distribution).

- Different scenarios for weighting were considered:
  - 1. Standardised weights
  - 2. Trimming of the weights

---


Lilly Confidential. Copyright © 2015. Internal Use Only. DO NOT COPY. Not for use in promotion. This is a forward-looking planning document. It is not our intention to implement marketing communications prior to regulatory approval of anticipated approved indications. Subject to approval, any promotional activities will be developed/implemented consistent with approved label and carried out in compliance with all company policies and applicable laws and regulations.
Case Study: Entropy balancing

Entropy balancing relies on a maximum entropy re-weighting scheme, which matches the empirical moments of the two data sources.

For this study constraints were applied to the first moments only. We reweight the RCT patients so that the covariate moments match those observed in the cohort study defining $x_{RCT,ik}$ (resp. $x_{COH,jk}$) as the value of the $k$-th covariate for patient $i$ in the RCT (resp. patient $j$ in the cohort study), and setting $N_{RCT}$ and $N_{COH}$ as the number of patients in the RCT and cohort respectively, the entropy weights $\omega = \{\omega_i, i = 1, ..., N_{RCT}\}$ satisfy the following optimization problem:

$$\min_\omega H(\omega) = \sum_{i=1}^{N_{RCT}} \omega_i \log(N_{RCT} \omega_i)$$

such that

$$\sum_{i=1}^{N_{RCT}} \omega_i (x_{RCT,ik})^r = \sum_{j=1}^{N_{COH}} (x_{COH,jk})^r \text{ for all } k, \text{ and for } r = 1, ..., R,$$

$$\sum_{i=1}^{N_{RCT}} \omega_i = N_{COH}, \text{ and } \omega_i \geq 0 \text{ for all } i = 1, ..., N_{RCT}.$$

The number $R$ represents the highest order of moments with respect to which we balance the two samples.

This optimization program above describes a specific case of the entropy balancing scheme.

For estimating treatment effect and confidence intervals, we replicate the approach for IPS.

Case Study: Methods

1. Calculate weights (using RWD)
2. Apply weights to outcome (RCT)
3. Record Weighted.TE(RCT)

Uncertainty estimated using bootstrapping:
1. Take 1000 samples; 
   \( n=n_{RCT} \) from RCT with replacement
2. Repeat the calculation of weights for each sample

Repeat 1000x

Error(HR)

LCL= 2.5 Percentile
UCL= 97.5 Percentile

TE= Median= 50th Percentile
Variables including in balance

♦ Variables selected based on availability in both data sets & some were considered effect modifiers
  • Age (70 years older vs younger 70 years)
  • Gender (female vs male)
  • Race (Asian vs non-Asian)
  • Smoking status (ex-smoker vs current smoker, never smoker vs current smoker)
  • Basis for diagnosis (histologic or cytological)
  • Time since diagnosis of NSCLC to study entry
  • Diagnosis subtype (adenocarcinoma vs large cell)
  • Stage of disease at study entry (IIIB vs IV)
  • ECOG performance status (0 vs 1)
  • Number of metastatic sites
  • Prior surgery (yes or no)
  • Prior radiotherapy (yes or no)
  • Presence of cardiovascular condition
  • Presence of lung conditions
  • Diabetes
Results: Baseline characteristics for included patients in JMDB (RCT) & FRAME (RWE)

<table>
<thead>
<tr>
<th></th>
<th>FRAME (N=948)</th>
<th>JMDB (N=1209)</th>
<th>p-Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years – mean (SD)</td>
<td>62.3 (9.86)</td>
<td>59.7 (9.34)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time since diagnosis – month (SD)</td>
<td>2.8 (12.61)</td>
<td>1.9 (7.75)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>293 (31%)</td>
<td>405 (33.5%)</td>
<td>0.211</td>
</tr>
<tr>
<td>Non-Asian n (%)</td>
<td>930 (98%)</td>
<td>997 (83%)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>297 (31%)</td>
<td>277 (23%)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>484 (51%)</td>
<td>585 (48%)</td>
<td></td>
</tr>
<tr>
<td>Never smoker</td>
<td>121 (13%)</td>
<td>195 (16%)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>46 (5%)</td>
<td>152 (13%)</td>
<td></td>
</tr>
<tr>
<td>Diagnosis n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytological</td>
<td>242 (26%)</td>
<td>453 (38%)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Histopathological</td>
<td>706 (74%)</td>
<td>756 (62%)</td>
<td></td>
</tr>
<tr>
<td>Diagnosis subtype n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>725 (77%)</td>
<td>861 (71%)</td>
<td>0.005</td>
</tr>
<tr>
<td>Large Cell Carcinoma</td>
<td>77 (8%)</td>
<td>145 (12%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>146 (15%)</td>
<td>203 (17%)</td>
<td></td>
</tr>
<tr>
<td>Stage n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIB</td>
<td>206 (22%)</td>
<td>272 (23%)</td>
<td>0.676</td>
</tr>
<tr>
<td>IV</td>
<td>742 (78%)</td>
<td>937 (77%)</td>
<td></td>
</tr>
<tr>
<td>ECOG n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>275 (29%)</td>
<td>446 (37%)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>1</td>
<td>673 (71%)</td>
<td>763 (63%)</td>
<td></td>
</tr>
<tr>
<td>Number of Metastatic Sites n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>771 (81%)</td>
<td>288 (24%)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>2</td>
<td>157 (17%)</td>
<td>296 (24%)</td>
<td></td>
</tr>
<tr>
<td>3+</td>
<td>20 (2%)</td>
<td>625 (52%)</td>
<td></td>
</tr>
<tr>
<td>Prior radiotherapy n (%)</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Prior Surgery n (%)</td>
<td>Yes</td>
<td>Yes</td>
<td>0.122</td>
</tr>
<tr>
<td>CV History n (%)</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lung History n (%)</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

- majority of patients in the FRAME study had 0-1 metastases compared to the JMDB study, which had sicker patients with more than half having 3+ metastases
Results: Standardised Difference Plot: JMDB (RCT) & FRAME (RWE)

- Standardized difference plot for differences between studies
- JMDB and FRAME – propensity score weighting
Results: Distributions of propensity scores* for JMDBG RCT (Red) and FRAME RWE study (Blue)

- While the distributions largely overlap, it is clear that the populations are substantially different – probably driven by the pronounced differences in number of metastases

*unstandardized and untrimmed weights for primary cohort (generated including all main effects
**Results:** Propensity weighted* overall survival re-analysis of JMDB RCT using FRAME RWE study for primary cohort

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Median Time to OS (months)</th>
<th>Hazard Ratio</th>
<th>95% LCL</th>
<th>95% UCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>No weighting*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>608</td>
<td>10.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>614</td>
<td>11.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted analysis: Unadjusted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>N</td>
<td>Median time to OS (months)</td>
<td>Hazard Ratio</td>
<td>Bootstrap 2.5 percentile</td>
<td>Bootstrap 97.5 percentile</td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>593</td>
<td>10.15</td>
<td>0.915</td>
<td>0.599</td>
<td>1.333</td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>616</td>
<td>15.57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*including all main effects - unstandardized untrimmed weights

**differences to Scagliotti et al (2008) are due to slightly different patient population included for analysis

≫ As reference JMDB** (RCT) results:
Results: Propensity weighted overall survival re-analysis of JMDB RCT using FRAME RWE study for primary cohort

- fairly consistent results for HR
- better balance comes at the expense of higher variability

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Median time to OS (months)</th>
<th>Hazard Ratio</th>
<th>Bootstrap 2.5 perc</th>
<th>Bootstrap 97.5 perc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propensity score weighting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>593</td>
<td>10.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>616</td>
<td>15.57</td>
<td>0.915</td>
<td>0.599</td>
<td>1.333</td>
</tr>
<tr>
<td>Propensity score weighting (standardized weights)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>593</td>
<td>10.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>616</td>
<td>14.09</td>
<td>0.878</td>
<td>0.705</td>
<td>1.086</td>
</tr>
<tr>
<td>Propensity score weighting (excluding number of metastasis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>593</td>
<td>9.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>616</td>
<td>11.24</td>
<td>0.797</td>
<td>0.622</td>
<td>1.20</td>
</tr>
<tr>
<td>Entropy balancing weighting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>593</td>
<td>10.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>616</td>
<td>14.95</td>
<td>0.904</td>
<td>0.602</td>
<td>1.293</td>
</tr>
<tr>
<td>Entropy balancing weighting (excluding number of metastasis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>593</td>
<td>9.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>616</td>
<td>11.24</td>
<td>0.789</td>
<td>0.629</td>
<td>1.004</td>
</tr>
</tbody>
</table>
Limitations

♦ Definitions of variables can be different between RCT and RWE studies
  • Baseline characteristics
  • Outcome measures
♦ Unmeasured confounders
♦ Non-overlapping propensity scores
♦ Extreme weights
♦ Specific categories of a variable are not available in RCT
Conclusions

♦ The key objective of this case study was to assess the generalizability of RCT results for the treatment of non-squamous NSCLC when projected to a real-world population.

♦ Tested reweighting efforts did not seem to invalidate findings from the original RCT.

♦ Representatives from Health Technology Assessment and Regulatory authorities as NICE, HAS, CADTH and EMA, as well as renowned academic institutions favourably commented on this approach with regard to impact on health care decision making.

♦ It can be used to answer the question,

"what would the trial results look like if they were conducted in a population that is reflective of the local population in which this new treatment is being launched into."
AIM: To develop a method to predict drug effectiveness prelaunch and to apply it in a case study in rheumatoid arthritis (RA).

The approach:
1. identifies a market-approved treatment (S) currently used in a target population similar to that of the new drug (N);
2. quantifies the impact of treatment, prognostic factors, and effect modifiers on clinical outcome;
3. determines the characteristics of patients likely to receive N in routine care;
4. predicts treatment outcome in simulated patients with these characteristics.

♦ Sources of evidence include expert opinion, RCTs, and observational studies.
♦ The framework relies on generalized linear models.
♦ Has not been tested on survival outcomes which presents new challenges
Challenges with Survival data

♦ How to specify the ‘correct’ model?
♦ How to simulate survival time?
♦ How to simulate censoring?
♦ What do to in the case of proportional hazards assumption being violated?
♦ How practical is this approach if you are unable to validate the chosen model in other data sources?
Questions?
## Case Study: NSCLC

Non-small cell lung cancer accounts for 85% of all lung cancer\(^1\)

<table>
<thead>
<tr>
<th>Non-small cell lung cancer histology subtypes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-squamous (76%)(^1)</strong></td>
<td><strong>Squamous (24%)(^1)</strong></td>
</tr>
<tr>
<td>Major subtype: adenocarcinoma(^1,2)</td>
<td>Subtypes: basaloid, clear cell type, papillary, small cell non-keratinizing(^8)</td>
</tr>
<tr>
<td>Less common subtypes: large cell, sarcomatoid, and adenosquamous carcinoma(^3)</td>
<td></td>
</tr>
<tr>
<td>Cells form recognizable glandular patterns(^4)</td>
<td>Characterized by high keratin production(^8)</td>
</tr>
<tr>
<td>Usually develops in lung periphery(^2,5)</td>
<td>Sheets or islands of large polygonal malignant cells and intercellular bridges(^8)</td>
</tr>
<tr>
<td>Relatively more common in women(^1)</td>
<td>Relatively more common in men(^1)</td>
</tr>
<tr>
<td>Affects smokers and non-smokers(^6)</td>
<td>Linked more strongly with smoking(^6)</td>
</tr>
<tr>
<td>Incidence increasing in recent years(^7)</td>
<td>Incidence decreasing in recent years(^7)</td>
</tr>
</tbody>
</table>


---

Non-small cell lung cancer (NSCLC) is the most common type of lung cancer, accounting for 85% of all lung cancer cases. It is further divided into non-squamous and squamous subtypes. The non-squamous subtype, which includes adenocarcinoma, is the most common, while squamous cell carcinoma, accounting for 24% of cases, is less common.

### Non-squamous NSCLC

- **Major subtype:** Adenocarcinoma
- **Less common subtypes:** Large cell, sarcomatoid, and adenosquamous carcinoma
- **Cells form:** Recognizable glandular patterns
- **Development:** Usually in the lung periphery
- **Common in:** Women
- **Affects:** Smokers and non-smokers
- **Incidence:** Increasing in recent years

### Squamous NSCLC

- **Subtypes:** Basaloid, clear cell type, papillary, small cell non-keratinizing
- **Characterized by:** High keratin production
- **Cells:** Sheets or islands of large polygonal malignant cells and intercellular bridges
- **Location:** Generally located centrally and in the larger bronchi
- **Common in:** Men
- **Linked to:** Strongly with smoking
- **Incidence:** Decreasing in recent years

---

### Notes