

# Causal Inference and Statistical Considerations for Indirectly Comparing Time-to-Event Endpoints for Treatments with Different Starting Points for Outcome Assessment in Resectable NSCLC

Belletti R<sup>1</sup>, Jovanoski N<sup>1</sup>, Abogunrin S<sup>1</sup>, Lee JS<sup>2</sup>, Crama L<sup>3</sup>, Abdul-Ghani R<sup>4</sup>, Gsteiger S<sup>1</sup>  
<sup>1</sup>F. Hoffmann-La Roche Ltd, Basel, BS, Switzerland; <sup>2</sup>Genentech, Inc., South San Francisco, CA, USA;  
<sup>3</sup>Roche Farma SA, Madrid, Spain; <sup>4</sup>Hoffmann-La Roche Limited, Mississauga, ON, Canada

## Background

Recently, pre-operative (neo-adjuvant), post-operative (adjuvant) and peri-operative (both neo-adjuvant and adjuvant) immunotherapy-based treatments have shown positive results in clinical trials for resectable non-small-cell lung cancer (NSCLC) (Provencio et al. 2022). Some of these therapies have also been approved by regulatory authorities, therefore different treatment options are or will soon be available for patients and there is an ongoing discussion regarding the optimal treatment sequence and its use and applicability in clinical practice (Mountzios et al. 2023). In the absence of clinical trials providing direct comparisons of these treatment strategies, indirect treatment comparisons can provide comparative effectiveness estimates. There has been a recent attempt to indirectly compare peri-operative with post-operative immunotherapy-based treatments (Goring S. et al. 2023). Given that these treatments have different starting points with respect to diagnosis and surgery, such comparisons are prone to bias. Additional methodological considerations are needed for a causal interpretation of the results of such analyses.

## Objectives

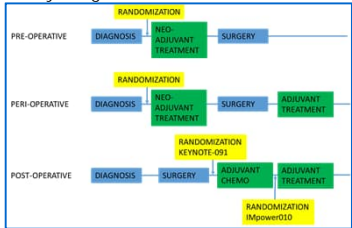
The objective of this research was to assess the conditions and assumptions needed to allow a causal interpretation of indirect treatment comparisons, when treatments start at different time points following diagnosis and time-to-event endpoints are considered. This evaluation can highlight potential biases if these conditions and assumptions are not met.

## Methods

We considered trials in resectable NSCLC with pre-, peri- and post-operative immunotherapy-based treatments, where outcomes are measured starting from different time points with respect to disease diagnosis and surgery. We analyzed the treatment decision problem and assessed its practical feasibility, the conditions for the transitivity assumption needed for network meta-analysis (NMA) and additional assumptions. Usual pre-requisites for the transitivity assumptions are (Chaimani et al. 2023, Cochrane handbook chapter 11):

- the clinical trials share the same target population
- the distribution of the effect modifiers is similar across the studies
- the interventions considered in the clinical trials are alternative options to be potentially chosen in practice by patients and health professionals and “jointly randomizable”, i.e. as if patients could be randomized to any of the treatment options in an hypothetical clinical trial
- the outcome definition and assessment is similar across the studies.

Figure 1 – Pre-, peri- and post-operative immunotherapy-based treatments: study design



## Results

**DECISION PROBLEM**  
The decision that patients and physicians face is which immunotherapy-based treatment strategy to adopt when patients are diagnosed with resectable NSCLC. This decision problem is only relevant to patients for whom all treatment strategies are equally applicable and it may be limited in clinical practice. Additional considerations are usually taken into account for the decision and these are not necessarily captured in clinical trials which do not include all treatment strategies. For example, adjuvant treatment may be preferable to allow timely surgical intervention to prevent early growth and spread of the disease; if accurate pathological staging did not occur before surgery, resection could allow for correct pathological staging, which has important prognostic considerations; starting systemic therapy after surgery, would allow enough time and tissue for biomarker testing; finally, not all patients would have the opportunity to receive neo-adjuvant therapy, as not all centers are currently able to integrate neo-adjuvant regimens into clinical practice. Examples of reasons for choosing neo-adjuvant treatments are to enhance systemic treatment compliance, the possibility of downstaging and an early assessment of systemic therapy impact (Provencio et al. 2022, Mountzios et al. 2023).

**TARGET POPULATION**  
The target population of the trials involving only one treatment modality (e.g. post-operative) could differ, as patients may be excluded, if the alternative modality (e.g. pre- or peri-operative) is more suitable based on the previously mentioned considerations. Therefore, some patient characteristics, not necessarily known or collected, may act as effect modifiers and bias the comparisons of treatments across trials. To emulate a clinical trial where patients could be randomized to any of these treatment strategies, the population characteristics of the different trials should be compared at the same time-point i.e. at the earliest treatment start after diagnosis as in pre- (e.g. CheckMate 816, Forde P.M. et al.) or peri- (AEGEAN, Heymach J. et al. 2023; Neotorch, Lu S. et al. 2023; KEYNOTE-671, Wakalee H et al. 2023) operative trials. In trials with post-operative treatments the population characteristics were assessed after surgery (e.g. in PEARLS/KEYNOTE-091 trial, O'Brien et al. 2022) or after adjuvant chemotherapy following surgery (e.g. in IMpower010, Felip E. et al. 2021). Therefore, it is impossible to assess, for the given set of trials, whether the actually enrolled patient populations are comparable in terms of effect modifiers. An illustration of the different treatment strategies employed in resectable NSCLC clinical trials is provided in Figure 1.

**INTERVENTIONS**  
Pre-, peri- and post-operative treatment strategies start at different time-points after diagnosis and relative to surgery. This implies that the randomization occurs at different time-points: shortly after the diagnosis for trials with pre- or peri-operative treatments, or after surgery or after adjuvant chemotherapy following surgery for trials with post-operative treatments (Figure 1).

**OUTCOME DEFINITION AND ASSESSMENT**  
The outcome definition and assessment is also different, if the treatment starting point is different. As outlined in Table 1, the outcome definition differs in particular between trials with pre-/peri-operative treatments, where event-free survival was used (progression before or after surgery, or recurrence after surgery) and trials with post-operative treatments, where disease-free survival was used (recurrence or new primary NSCLC). The outcome is assessed from randomization which has a different timing in trials with pre- and peri- compared to trials with post-operative treatments. An illustration of the different outcome assessment and definitions in trials in resectable NSCLC is provided in Figure 2.

**ANALYTICAL METHODS**  
If the assumptions illustrated in the «Methods» section hold, then we can assume transitivity of the relative effect and perform an NMA with methods for continuous endpoints. Additional assumptions are the log-normal distribution for the HR and the proportional hazards. When comparing pre- and peri- to post-operative treatments, it appears that the assumptions either are not fulfilled or are not evaluable. Analytical strategies aimed at estimating and combining time-varying HRs to account for the different treatment starting points as used by Goring S. et al 2023 are not a solution, as time varying HRs are inherently biased (Hernan M. 2010), since they are conditional on the probability of surviving until a specific time which may depend on the treatment received by and baseline characteristics of patients. The different treatment timing results in patient attrition cause selection bias, which is difficult to account for with analytical methods, if individual patient data is not available for all studies.

Figure 2 – Pre-, peri- and post-operative immunotherapy-based treatments: duration, outcome definition and assessment

Study type	Study	Neo-adjuvant treatment	Neo-adjuvant treatment duration	% pts with surgery	Adjuvant treatment	Adjuvant treatment duration	Study type	Study	Outcome	Definition
Pre-operative	CHECKMATE-816	Nilv + chemo	≤ 9 weeks	83.2	NA		Pre-operative	CHECKMATE-816	EFS	(BICR) any progression of disease precluding surgery, progression or recurrence of disease after surgery, progression of disease in the absence of surgery, or death from any cause
Pre-operative	AEGEAN	Durva + chemo	≤ 12 weeks	80.6	Durva	≤ 48 weeks	Pre-operative	AEGEAN	EFS	(A) progressive disease (PD) that precludes surgery, (B) PD discovered and reported by the investigator upon attempting surgery that prevents completion of surgery; (C) local/distant recurrence using BICR per RECIST v1.1; or (D) death from any cause
Pre-operative	KEYNOTE-671	Pembro + chemo	≤ 12 weeks	82.1	Pembro	≤ 39 weeks	Pre-operative	KEYNOTE-671	EFS	First occurrence of local progression that precluded the planned surgery, unresectable tumor, progression or recurrence according to the RECIST v1.1, by the investigator's assessment, or death from any cause
Post-operative	IMpower010	Yes	3-12 weeks	Atzo BSC	Atzo BSC	≤ 48 weeks	Post-operative	IMpower010	DFS	Investigator-assessed first NSCLC recurrence, occurrence of new primary NSCLC, or death from any cause, whichever occurred first
Post-operative	PEARLS/KEYNOTE-091	Optional (86%)	≤ 12 weeks	Pembro Placebo	Pembro Placebo	≤ 54 weeks	Post-operative	PEARLS/KEYNOTE-091	DFS	Locoregional or metastatic recurrence assessed per RECIST version 1.1 by investigator review, appearance of a second NSCLC primary or other malignancy, or death from any cause, whichever occurred first

## Conclusions

Indirect treatment comparisons in resectable NSCLC comparing pre- or peri- to post-operative immunotherapies are methodologically challenging for several reasons: it is unclear if the clinical trial populations can be considered exchangeable, the outcome definition and assessment differs, as the randomization and the starting point of the treatment is different resulting in different patient attrition. Hence, it is unlikely that transitivity holds to employ traditional NMA methods for continuous endpoints assuming proportional hazards. New methods combining time-varying HRs do not offer a solution, if assumptions are not fulfilled and do not account for the causes of the selection bias. Alternative methods would need to account for effect modifiers and the factors causing patient attrition which otherwise induces selection bias.