



Brief report

Methodological interpretation of subgroup analysis by histological subtype for perioperative toripalimab in resectable non-small-cell lung cancer

Manuel David Gil-Sierra*, Elisa Pizarro-Barron and Maria del Pilar Briceño-Casado

Servicio de Farmacia, Hospital Universitario de Puerto Real, Puerto Real, Spain

ARTICLE INFO

Article history:

Received 6 May 2025

Accepted 27 October 2025

Available online xxxx

Keywords:

Toripalimab

Lung cancer

Comparative histology

Treatment outcome

Statistical interpretation of data

A B S T R A C T

Objective: Pivotal study on perioperative toripalimab in resectable non-small-cell lung cancer showed population differences by histological subtype compared with other immunotherapy regimens, raising doubts about therapeutic positioning. The aim of this study was to interpret the methodological analysis by subgroups according to histological subtype of perioperative toripalimab in resectable non-small-cell lung cancer.

Methods: Validated subgroup analysis applicability tool was used. This tool had two parts: preliminary questions to directly rule out analysis without relevant minimum conditions, and checklist. This checklist assessed statistical association, biological plausibility and consistency of subgroup results, and related these criteria to recommendations on applicability.

Results: Preliminary question regarding differences in effect between subgroups $p(i) < 0.1$ was answered negatively, and checklist was not applied due to direct discard. Even if the checklist had been applied, statistical association criterion would have been rated 'null' due to absence of statistically significant differences. Biological plausibility would have been rated 'probable' due to non-squamous histology being a negative prognostic factor. Consistency would have been rated 'null' for absence of heterogeneity between subgroups in similar studies.

Conclusions: This methodological interpretation recommended against applying histology-based subgroup results for perioperative toripalimab in resectable non-small-cell lung cancer, avoiding ruling out the use of toripalimab in the non-squamous subgroup.

© 2025 Sociedad Española de Farmacia Hospitalaria (S.E.F.H). Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Interpretación metodológica de análisis por subgrupos según el subtipo histológico para toripalimab perioperatorio en cáncer de pulmón no microcítico reseccable

R E S U M E N

Objetivo: el estudio pivotal sobre toripalimab perioperatorio en cáncer de pulmón no microcítico reseccable presentó diferencias poblacionales según el subtipo histológico frente a los demás esquemas inmunoterápicos, generando dudas sobre su posicionamiento terapéutico. El objetivo es realizar una interpretación metodológica del análisis por subgrupos según el subtipo histológico de toripalimab perioperatorio en cáncer de pulmón no microcítico reseccable.

Método: se utilizó una herramienta validada sobre aplicabilidad del análisis por subgrupos con 2 partes: cuestiones preliminares para descartar directamente los análisis sin condiciones mínimas relevantes y un *checklist*. Este *checklist* valoró la asociación estadística, la plausibilidad biológica y la consistencia de los resultados de los subgrupos, relacionando estos criterios con recomendaciones sobre su aplicabilidad.

Resultados: la cuestión preliminar sobre diferencias de efecto entre subgrupos ($p(i) < 0,1$) tuvo respuesta negativa, por lo que el *checklist* no fue aplicado debido al descarte directo. Aunque el *checklist* hubiera sido aplicado, el criterio de asociación estadística habría obtenido una valoración «nula» por ausencia de diferencias estadísticamente significativas. La plausibilidad biológica habría alcanzado una valoración «probable» por histología no escamosa como factor pronóstico negativo. La consistencia hubiera sido valorada como «nula»

Palabras clave:

Toripalimab

Cáncer de pulmón

Histología comparativa

Resultado de tratamiento

Interpretación estadística de datos

DOI of original article: <https://doi.org/10.1016/j.farma.2025.10.018>.

* Corresponding author.

E-mail address: mangilsie@yahoo.com (M.D. Gil-Sierra).

<https://doi.org/10.1016/j.farma.2026.02.017>

1130-6343/© 2025 Sociedad Española de Farmacia Hospitalaria (S.E.F.H). Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

por la ausencia de heterogeneidad entre subgrupos en estudios similares.

Conclusiones: esta interpretación metodológica recomendó no aplicar los resultados de los subgrupos según histología para toripalimab perioperatorio en cáncer de pulmón no microcítico reseccable, evitando descartar su uso en el subgrupo no escamoso.

© 2025 Sociedad Española de Farmacia Hospitalaria (S.E.F.H). Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Adjuvant chemotherapy after surgical resection has long been the standard of care for resectable non-small-cell lung cancer (NSCLC).¹ In this setting, the emergence of immunotherapy marked a significant advance. For example, the use of adjuvant pembrolizumab was shown to increase disease-free survival in patients with resectable NSCLC at stages IB–IIIA.²

Subsequently, the use of perioperative immunotherapy in resectable NSCLC was evaluated. The perioperative approach involves using drugs both before and after surgery. Several randomised phase III clinical trials have analysed the results of immune checkpoint inhibitor regimens in this clinical setting.^{3–6} These regimens offer notable benefits, as they are associated with increased event-free survival.

Studies are needed that indirectly compare the different perioperative immunotherapy regimens in order to clarify their relative therapeutic position in resectable NSCLC.⁷ In addition, analyses are needed to determine whether these clinical trials have comparable characteristics and populations or whether such indirect comparisons should be rejected.⁸ Evaluation of the different therapies must take into account biomarkers, including programmed death-ligand 1 (PD-L1), as well as patient characteristics such as lymph node count and disease extent. The efficacy of a given treatment may differ from that of others, for example, due to different PD-L1 expression levels or disease stage.

The Neotorch clinical trial evaluated the perioperative use of toripalimab in resectable NSCLC.⁶ The patient population in this clinical trial differed from those enrolled in studies of other perioperative immunotherapy regimens. The trial found that 78% of patients had squamous histology, compared with 43% to 50% in other studies. Given that histological subtypes can influence disease prognosis, this difference could be considered a bias when comparing different perioperative immunotherapy combinations in resectable NSCLC.⁹ Further analysis of the results according to histological subtype in this clinical setting could help to determine the most appropriate therapeutic approach. This would prevent the misinterpretation of subgroup results from adversely influencing clinical decision-making.

Subgroup analysis evaluates data obtained from a healthcare intervention across different subpopulations according to a given factor.¹⁰ This approach must be applied with caution, as it inherently increases uncertainty.¹¹ On the one hand, additional determinations increase the likelihood of detecting differences that are not truly present (Type I error). On the other hand, the redistribution of patients into smaller subgroups can prevent the detection of real differences (Type II error). Therefore, subgroup analyses should be conducted systematically before making clinical decisions or determining therapeutic approaches.

The objective of this study was to methodologically interpret subgroup analysis by histological subtype on the perioperative use of toripalimab in resectable NSCLC in the Neotorch clinical trial.

Methods

Using a validated tool developed by Gil-Sierra et al., the researchers systematically interpreted subgroup analyses by histological subtype, enhancing the clinical applicability of the results.¹² The tool comprises 2 parts: 4 preliminary questions to determine whether the subgroup

analysis meets the minimum relevant conditions, followed by a checklist.

The preliminary questions address the level of evidence of the study with subgroup analysis, the clinical relevance of the variable evaluated, the difference in effect between subgroups, and the presence of the factor determining the analysis prior to the healthcare intervention. Progression to the checklist requires that all preliminary questions be satisfied: failure to satisfy any of them results in exclusion of the subgroup analysis. This approach provides a rapid way to prevent conducting a subgroup analysis that does not meet the minimum relevant conditions.

The second part of the tool (the checklist) is used to evaluate a set of criteria, including the statistical evidence supporting subgroup effects, the biological plausibility of differences between subgroups, and the consistency of findings with similar studies. The criterion relating to statistical evidence was further subdivided into the assessment of the interaction p-value to estimate the probability that observed differences between subgroups are attributable to chance, taking into account prespecification of the subgroup analysis, sample size, the number of factors evaluated, and the overall study outcome. Where data were not provided, the p-value was quantified using the subgroup calculator described by Primo et al.¹³

For each criterion, the checklist was used to assign an association score reflecting the strength of evidence: probable (+3 points), possible (+2), doubtful (0), and none (–3). The tool itself includes a guide containing clarifications and comments to facilitate the assignment of scores when evaluating criteria. The total score can be used to guide recommendations on the use of subgroup data in clinical decision-making. Higher scores indicate greater reliability: probable (7–9 points), supporting the application of subgroup results until another confirmatory randomised clinical trial is published; possible (5–6 points), allowing cautious use due to factors such as low drug tolerance, difficulty of administration, or cost; doubtful (3–4 points), generally precluding application except in exceptional cases; and nil (<3 points), indicating that the subgroup analysis should not be applied.

Results

Of the 4 preliminary questions on minimum conditions, 3 were satisfied. Subgroup analysis was conducted in the Neotorch phase III randomised clinical trial. Phase III trials of this type provide the highest level of scientific evidence. We selected event-free survival as the variable for the subgroup analysis due to its clinical relevance. This variable was defined as the time from randomisation to disease progression precluding surgery, postoperative progression, local or metastatic recurrence, or death. The histological subtype evaluated in the subgroup analysis was determined before drug administration.

The preliminary test for differences in treatment effect between subgroups did not demonstrate a statistically significant interaction at the 0.1 level. In the Neotorch trial, the interaction p-value between subgroups defined by histology was 0.29.⁶ Therefore, the checklist was not applied due to direct exclusion. The statistical association criterion would have obtained a nil score (–3 points) due to the absence of a statistically significant difference between the squamous and non-squamous histology subgroups. The criterion of biological plausibility would have received a probable score (+3 points), given that a

Table 1
 Interaction p-value between subgroups by histological subtype based on efficacy values for event-free survival during the consistency assessment.

Study (author and year)	Immunotherapeutic agent used in the perioperative setting	EFS value in the squamous histology subgroup (HR and CI)	EFS value in the non-squamous histology subgroup (HR and CI)	Interaction p-value
Heymach et al. (2023)	Durvalumab	0.71 (95%CI: 0.49–1.03)	0.69 (95%CI: 0.48–0.99)	0.91
Wakelee et al. (2023)	Pembrolizumab	0.57 (95%CI: 0.41–0.77)	0.58 (95%CI: 0.43–0.78)	0.94
Cascone et al. (2024)	Nivolumab	0.46 (95%CI: 0.30–0.72)	0.72 (95%CI: 0.49–1.07)	0.13
Lu et al. (2024)	Toripalimab	0.35 (95%CI: 0.23–0.52)	0.54 (95%CI: 0.26–1.08)	0.29

CI, confidence interval; EFS, event-free survival; HR, hazard ratio.

hypothesis supported by previous literature could explain the observation. Previous studies have suggested that non-squamous histology is a factor for poor prognosis.⁹ This finding could be consistent with the hazard ratio (HR) and confidence interval (CI) for the non-squamous histology subgroup in the Neotorch trial (HR = 0.54, 95%CI: 0.26–1.08), which includes the null value.⁶ However, the consistency criterion would have been rated as nil (– 3 points). Previous studies evaluating the use of perioperative immunotherapy found no statistically significant difference in p-values between squamous and non-squamous histology subgroups. Regarding event-free survival, the interaction p-value was 0.94 for perioperative pembrolizumab, 0.13 for nivolumab, and 0.91 for durvalumab (Table 1). These findings support rejecting the assumption of differences between the 2 subgroups, despite numerically larger hazard ratios observed in the non-squamous histology subgroups for perioperative toripalimab, pembrolizumab, and nivolumab. Table 2 summarises the interpretation of the subgroup analysis by histological subtype in the Neotorch study.

Discussion

This methodological interpretation suggests that histology-based subgroup results for perioperative toripalimab in resectable NSCLC should not be used in place of the overall results. This conclusion was reached through a systematic methodological assessment of the results from the squamous and non-squamous histology populations in the Neotorch trial, contextualised with the findings of other studies of perioperative immunotherapies in this setting.^{3–6,12}

In the squamous subgroup, toripalimab showed a confidence interval that did not cross the null value and a lower numerical hazard ratio compared with the non-squamous subgroup. However, no interaction between histological subtypes was found in the Neotorch study.⁶ Perioperative use of pembrolizumab and nivolumab in resectable NSCLC also showed larger numerical effects in the squamous subgroup than in the non-squamous subgroup, but these differences did not reach statistical significance. In addition, durvalumab showed a pattern of subgroup effects opposite to that observed in previous studies, although the differences between subgroups did not reach statistical significance. Therefore, the consistent lack of heterogeneity between these subgroups for immunotherapies with similar mechanisms of action and

clinical settings suggests that histology-based subgroup analysis should not be applied, based on current evidence. In addition, it should be noted that the low representation of patients with non-squamous histology in each arm (N = 45) of the Neotorch study further limits the applicability of subgroup analysis.

The main strength of the present study is the methodological assessment of subgroup analysis, which avoids interpretations based on unjustified a priori assumptions. Another strength of our study is that systematic evaluation of subgroup analyses in clinical trials is uncommon in the literature. A limitation inherent to subgroup results is the lack of conclusive findings from the clinical trials on which our research is based. In addition, the perioperative immunotherapy regimen used in the Neotorch trial⁶ was as follows: 3 neoadjuvant cycles of toripalimab plus chemotherapy prior to surgery, followed by an additional adjuvant cycle of toripalimab with chemotherapy, followed by toripalimab monotherapy. In contrast, the other perioperative immunotherapy regimens included 4 neoadjuvant cycles of immunotherapy plus chemotherapy prior to surgery, followed by the immunotherapy agent as adjuvant monotherapy.

The subgroup analysis approach requires the systematic interpretation of a set of relevant statistical and clinical criteria. Previous literature recommends which subgroup characteristics should be evaluated, such as a plausible explanation for the observed effect of the healthcare intervention in specific subpopulations or an appropriate temporal sequence between the factor studied and the drug administered.^{10,12} In fact, previous literature has emphasised the lack of credibility in conducting subgroup analyses based on irrelevant characteristics, such as zodiac sign. It should be remembered that the exploratory nature of subgroup analyses can generate numerous hypotheses that lack a sound pathophysiological basis.

Inadequate interpretation of subgroup analysis by histological subtype for perioperative treatment in resectable NSCLC could negatively affect therapeutic decisions, leading to consequences similar to those observed in other subgroup analyses in oncohaematology.¹⁴ In this context, the use of toripalimab in the non-squamous subgroup could be mistakenly ruled out. This situation would reduce the number of therapeutic alternatives for perioperative treatment of resectable NSCLC. Consequently, price competition would be reduced, making it difficult to acquire drugs at lower prices and to make optimal therapeutic

Table 2
 Summarised interpretation of the subgroup analysis by histological subtype in the Neotorch study.

Methodology	Criteria	Variable		
		Event-free survival		
Validated tool (Gil-Sierra et al.) ¹²	Preliminary questions	Highest level of evidence in the study with subgroup analysis	Yes	
		Clear clinical relevance of the variable analysed or of the primary surrogate variable	Yes	
		Difference in effect between subgroups for the factor analysed	No	
		p < .1	Yes	
	Checklist	Factor determining the subgroup analysis prior to healthcare intervention	Not applied	Yes
			Statistical association (score)	Not applied
			Biological plausibility (score)	None (– 3 points)
			Consistency (score)	Probable (+ 3 points)
			Recommendation for application (total score)	None (– 3 points)
				'None' due to direct exclusion from preliminary questions: 'do not consider subgroups'

decisions. Therefore, our findings may help prevent the selective application of histology-based subgroup results, influenced by particular interests, in perioperative immunotherapy for resectable NSCLC.

Contribution to the scientific literature

A methodological analysis of histology-based subgroups for perioperative immunotherapy in resectable non-small-cell lung cancer. The results may have implications for the therapeutic positioning of perioperative immunotherapy.

Ethical responsibilities

All the necessary ethical responsibilities regarding authorship and redundant publication were fulfilled. Ethical approval regarding the protection of human and animal subjects, as well as informed consent, was not required due to the study design.

CRedit authorship contribution statement

Manuel David Gil-Sierra: Writing – review & editing, Writing – original draft, Visualisation, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Elisa Pizarro-Barron:** Writing – review & editing, Writing – original draft, Visualisation, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Maria del Pilar Briceño-Casado:** Writing – review & editing, Writing – original draft, Visualisation, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

None declared.

Conflict of interest

None declared.

References

1. Pignon JP, Tribodet H, Scagliotti GV, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *J Clin Oncol*. 2008;26(21):3552–3559. doi:10.1200/JCO.2007.13.9030.
2. O'Brien M, Paz-Ares L, Marreaud S, et al. Pembrolizumab versus placebo as adjuvant therapy for completely resected stage IB-IIIa non-small-cell lung cancer (PEARLS/KEYNOTE-091): an interim analysis of a randomised, triple-blind, phase 3 trial. *Lancet Oncol*. 2022;23(10):1274–1286. doi:10.1016/S1470-2045(22)00518-6.

3. Wakelee H, Liberman M, Kato T, et al. Perioperative pembrolizumab for early-stage non-small-cell lung cancer. *N Engl J Med*. 2023;389(6):491–503. doi:10.1056/NEJMoa2302983.
4. Heymach JV, Harpole D, Mitsudomi T, et al. Perioperative durvalumab for resectable non-small-cell lung cancer. *N Engl J Med*. 2023;389(18):1672–1684. doi:10.1056/NEJMoa2304875.
5. Cascone T, Awad MM, Spicer JD, et al. Perioperative nivolumab in resectable lung cancer. *N Engl J Med*. 2024;390(19):1756–1769. doi:10.1056/NEJMoa2311926.
6. Lu S, Zhang W, Wu L, et al. Perioperative toripalimab plus chemotherapy for patients with resectable non-small cell lung cancer: the Neotorch randomized clinical trial. *JAMA*. 2024;331(3):201–211. doi:10.1001/jama.2023.24735.
7. Mei T, Zhou Q, Gong Y. Comparison of the efficacy and safety of perioperative immunochemotherapeutic strategies for resectable non-small cell lung cancer: a systematic review and network meta-analysis. *Clin Oncol (R Coll Radiol)*. 2024;36(2):107–118. doi:10.1016/j.clon.2023.12.006.
8. Ortega A, Fraga MD, Alegre-del-Rey EJ, et al. A checklist for critical appraisal of indirect comparisons. *Int J Clin Pract*. 2014;68(10):1181–1189. doi:10.1111/ijcp.12487.
9. Khan OA, Fitzgerald JJ, Field ML, et al. Histological determinants of survival in completely resected T1-2N1M0 non small cell cancer of the lung. *Ann Thorac Surg*. 2004;77(4):1173–1178. doi:10.1016/j.athoracsur.2003.08.080.
10. Sun X, Ioannidis JPA, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. *JAMA*. 2014;311(4):405–411. doi:10.1001/jama.2013.285063.
11. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technol Assess*. 2001;5(33):1–56. doi:10.3310/hta5330.
12. Gil-Sierra MD, Fénix-Caballero S, Abdel Kader-Martin L, et al. Checklist for clinical applicability of subgroup analysis. *J Clin Pharm Ther*. 2020;45(3):530–538. doi:10.1111/jcpt.13102.
13. Primo J, Escrig J. Metasurv: calculadora Excel para metaanálisis de supervivencia; 2008 [accessed 01 May 2025]. Available from: https://www.google.es/search?q=primo+j+metasurv+redcaspe&sca_esv=a4c5c968ead58a07&ei=nk6NZ9OBCMSgigPzLWRyAk&ved=0ahUKEwjTyPHmvoKLaxVE0AIHHcxajBjkQ4dUDCBA&uact=5&oq=primo+j+metasurv+redcaspe&gs_lp=Egxnd3Mtd2l6LXNlcnAiGXByaW1vIgo_gbWV0YXN1cnYgcmVky2FzcGUyBRAhGKABMgUQJRigAUjPHVAAWP4ccAB4AJABAjgBrAKgAbEOqgEHMC44LjEuMbgBA8gBAPgBAZgCCqAC4w7CAggQABiABBiiBMICBxAhGKABGAqYAwCSBwcvLjguMS4xoAffIA&scient=gws-wiz-serp
14. Gil-Sierra MD, Briceño-Casado MDP, Sánchez-Hidalgo M. Clinical and pharmacoeconomic impact of subgroup analysis in onco-hematological patients. *Support Care Cancer*. 2022;30(5):3761–3772. doi:10.1007/s00520-022-06823-y.