The 10th Conference on Mathematics in Engineering: Theory and Applications
Novi Sad, December 12-14th 2025

META-ANALYSIS OF IMMUNOTHERAPY OUTCOMES IN METASTATIC MELANOMA

Katarina Vidojević[1], Jovana Dedeić[2], Srđan Milićević[3], Jelena Ivetić[4], and Marija Delić[5]

https://doi.org/10.24867/META.2025.19 Article type: Review paper

Abstract. Meta-analytic methods are essential for quantitative evidence synthesis, offering a statistical framework for model comparison and heterogeneity assessment. This paper revisits our previous work through a concise methodological review and model-comparison meta-analysis, illustrating adaptive model selection for binary outcomes. Thirteen studies (>6,000 patients) comparing monotherapy and combination regimens in metastatic melanoma immunotherapy were analyzed. Binary outcomes (benefit and risk rates) were pooled using fixed-effects (FEM), random-effects (REM), and generalized linear mixed models (GLMM), depending on heterogeneity. The results highlight how model choice influences pooled estimates and inference reliability, underscoring the importance of flexible statistical modeling in complex datasets.

AMS Mathematics Subject Classification (2020): 62P10, 62F15.

Key words and phrases: meta-analysis; FEM; REM; GLMM; immune checkpoint inhibitors (ICIs); melanoma; benefits; adverse events; risks.

1. Introduction

Meta-analysis is a key statistical framework for quantitative synthesis in medical and technical sciences, particularly valuable in studies of rare conditions with limited statistical power. Beyond summarizing results, it enables formal model comparison and heterogeneity assessment through likelihood-based methods. In modern oncology, where immune checkpoint inhibitors yield complex and variable outcomes, such modeling ensures reliable inference. We present a compact model-comparison meta-analysis that adaptively selects among FEM, REM, and GLMM according to data variability. The analysis includes 13 studies on metastatic melanoma immunotherapy, comprising 22

[1]Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, e-mail: vidojevic9@uns.ac.rs
[2]Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, e-mail: radenovicj@uns.ac.rs
[3]Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, e-mail: srdjan88@uns.ac.rs
[4]Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, e-mail: jelenaivetic@uns.ac.rs
[5]Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, e-mail: marijadelic@uns.ac.rs

study arms and over 6,000 patients ([4]-[16]) and summarizes our previous work published in [1, 2, 3].

2. Methods

2.1. Statistical models

For binary outcomes we applied three modeling strategies depending on heterogeneity (Cochran's Q, I^2): (i) FEM when $I^2 < 25\%$, (ii) REM for moderate heterogeneity, and (iii) GLMM with logit link for extreme heterogeneity ($I^2 \geq 90\%$). The GLMM was fitted on logit-transformed proportions with a random intercept per study arm; inverse-logit transformations produced pooled proportions and confidence intervals (CIs). Group comparisons used two-proportion Z-tests on pooled estimates when appropriate.

2.2. Quantifying Heterogeneity

Heterogeneity across studies was evaluated using Cochran's Q statistic and the I^2 index. Cochran's Q tests whether observed variability among study outcomes exceeds that expected by chance alone: $Q = \sum_{i=1}^{k} w_i(p_i - \bar{p})^2$, where k denotes the number of study arms, p_i represents the observed event proportion in the i-th arm study, $\bar{p}$ is weighted pooled estimate of the event proportion and $w_i = \frac{1}{\sigma_i^2}$ is the inverse-variance weight, with $\sigma_i^2 = \frac{p_i(1-p_i)}{n_i}$, where n_i is the study sample size.

To interpret heterogeneity in percentage form, the I^2 statistic was calculated as:

$$I^2 = \max\left(0, \frac{Q - (k-1)}{Q}\right) \times 100\%.$$

The I^2 value represents the proportion of total variation in observed outcomes attributable to real differences between studies rather than random sampling error. In accordance with Cochrane guidelines, heterogeneity levels were classified as low ($I^2 < 25\%$), moderate ($25\% \leq I^2 < 90\%$), and substantial ($I^2 \geq 90\%$).

These thresholds determined the choice among three modeling frameworks: FEM, REM, and GLMM.

2.3. Fixed-Effects Model

Under the FEM, it is assumed that all studies estimate a common underlying true effect, and that any observed variation arises purely from within-study error. The pooled proportion estimate under the FEM is given by:

$$\bar{p}_{FEM} = \frac{\sum_{i=1}^{k} w_i p_i}{\sum_{i=1}^{k} w_i},$$

where the weights w_i correspond to the inverse of the within-study variance. The variance and standard error of the pooled estimate are:

$$\mathrm{Var}(\bar{p}_{FEM}) = \frac{1}{\sum_{i=1}^{k} w_i}, \qquad SE_{FEM} = \sqrt{\mathrm{Var}(\bar{p}_{FEM})}.$$

The corresponding 95% confidence interval is expressed as:

$$(\bar{p}_{FEM} - 1.96 \cdot SE_{FEM},\ \bar{p}_{FEM} + 1.96 \cdot SE_{FEM}).$$

This model was applied to subgroups exhibiting low between-study heterogeneity ($I^2 < 25\%$), yielding narrow confidence intervals that reflect the homogeneity of treatment effects.

2.4. Random-Effects Model

When moderate heterogeneity was detected ($25\% \leq I^2 < 90\%$), the REM was employed to account for both within- and between-study variability. Between-study variance (τ^2) was estimated as:

$$\tau^2 = \frac{Q - (k-1)}{C}, \quad \text{where} \quad C = \sum_{i=1}^{k} w_i - \frac{\left(\sum_{i=1}^{k} w_i^2\right)}{\sum_{i=1}^{k} w_i}.$$

The adjusted weights under the REM were computed as $w_i^* = \frac{1}{\sigma_i^2 + \tau^2}$. The pooled estimate under the REM is given by:

$$\bar{p}_{REM} = \frac{\sum_{i=1}^{k} w_i^* p_i}{\sum_{i=1}^{k} w_i^*}.$$

This approach provided wider confidence intervals than FEM, capturing uncertainty arising from cross-study differences such as patient demographics, dosing schedules, and study design variability.

2.5. Generalized Linear Mixed Model

For datasets characterized by extreme heterogeneity ($I^2 \geq 90\%$), a GLMM was adopted. The GLMM extends traditional meta-analytic models by incorporating random effects directly into the link function. Event proportions were transformed via the logit function $y_i = \ln\left(\frac{p_i}{1-p_i}\right)$, and modeled as:

$$y_i = \beta_0 + u_i,$$

where β_0 represents the fixed intercept and $u_i \sim \mathcal{N}(0, \tau^2)$ denotes a random intercept capturing between-study variability. The fitted estimates were transformed back to the proportion scale using the inverse-logit function.

The GLMM provided the most flexible framework, yielding stable and interpretable results even under non-normal or highly dispersed data.

3. Results

Across all datasets, treatment outcomes were reported in accordance with the RECIST criteria, classifying therapeutic responses into complete response (CR), partial response (PR), overall response (OR) and stable disease (SD). In addition to survival outcomes, the analysis evaluated the incidence of irAEs, including pneumonitis, colitis, diarrhea, and thyroid-related disorders (hyperthyroidism and hypothyroidism). The dual focus on both treatment efficacy and safety allowed for an integrated risk-benefit assessment of immunotherapy regimens.

The pooled incidence rates of irAEs for monotherapy and combination therapy are summarized in Table 1. The results indicate a markedly higher frequency of adverse events in the combination therapy group, where 93.2% of patients experienced at least one irAE compared to 81.9% in the monotherapy group. For instance, pneumonitis occurred in approximately 5.55% of patients receiving combination therapy versus 1.95% for monotherapy; hypothyroidism was reported in 18.78% versus 10.41%, and hyperthyroidism in 14.09% versus 3.39%.

Table 1: Pooled prevalence estimates of adverse events by therapy group, with heterogeneity indicators.

Adverse event	Therapy type	Study arms	Patients	I^2 (%)†	Point estimate	95% CI
At least one AE*	M	10	3222	99.53[g]	0.8194	(0.6562, 0.9152)
	C	6	1801	61.50[r]	0.9322	(0.9111, 0.9534)
Pneumonitis*	M	10	2650	31.99[r]	0.0195	(0.0126, 0.0265)
	C	6	1801	70.88[r]	0.0555	(0.0335, 0.0775)
Hypothyroidism*	M	10	2638	85.43[r]	0.1041	(0.0729, 0.1353)
	C	6	1801	47.72[r]	0.1878	(0.1597, 0.2160)
Hyperthyroidism*	M	7	2244	92.31[g]	0.0339	(0.0173, 0.0654)
	C	5	1707	81.59[r]	0.1409	(0.0983, 0.1836)
Colitis*	M	10	2558	80.44[r]	0.0300	(0.0161, 0.0439)
	C	6	1801	81.22[r]	0.0753	(0.0450, 0.1056)
Diarrhea*	M	9	2567	90.14[g]	0.1845	(0.1367, 0.2443)
	C	6	1801	88.67[r]	0.3457	(0.2734, 0.4180)
Increased AST*	M	3	1231	0.00[f]	0.0670	(0.0531, 0.0810)
	C	5	1487	89.74[r]	0.1732	(0.1078, 0.2385)
Increased ALT*	M	3	1231	0.00[f]	0.0738	(0.0592, 0.0884)
	C	5	1487	88.64[r]	0.1881	(0.1226, 0.2536)

Table 2: Comparison of clinical benefit outcomes between monotherapy and combination therapy, with pooled prevalence and heterogeneity indicators.

Outcome	Therapy type	Study arms	Patients	I^2 (%)†	Point estimate	95% CI
CR*	M	7	2458	91.46[g]	0.1251	(0.0858, 0.1791)
	C	5	885	59.29[r]	0.1749	(0.1348, 0.2151)
PR*	M	8	2504	79.49[r]	0.2029	(0.1663, 0.2394)
	C	5	885	0.00[f]	0.3448	(0.3135, 0.3761)
OR*	M	8	2504	94.31[g]	0.3156	(0.2391, 0.4037)
	C	5	885	64.94[r]	0.5224	(0.4652, 0.5796)
SD*	M	8	2504	91.73[g]	0.1676	(0.1257, 0.2199)
	C	5	885	78.09[r]	0.1007	(0.0589, 0.1425)

M - monotherapy; C - combination therapy

† Superscript indicates model type used: [f] FEM, [r] REM, [g] GLMM.

* Statistically significant difference between mono and combination groups at $p < 0.001$.

The pooled therapeutic response outcomes, including CR, PR, OR, and SD rates for both therapy groups are presented in Table 2. The combination therapy demonstrated a complete response rate of approximately 17.49%, compared to 12.51% for monotherapy. The partial response rate was 34.48% for the combination group versus 20.29% for monotherapy, resulting in an overall response rate of 52.24% versus 31.56%. Conversely, the rate of stable disease was slightly higher with monotherapy (16.76%), suggesting that tumor regression was less frequent compared to the combination approach. These results highlight that combination immunotherapy yields substantially improved clinical efficacy but at the cost of an increased incidence of adverse events.

4. Conclusion

The results confirm that combination immune checkpoint blockade offers greater benefit than monotherapy but with higher immune-related toxicity, requiring careful monitoring and individualized treatment. Methodologically, integrating FEM, REM, and GLMM highlights the need to account for heterogeneity for accurate effect estimation. Flexible models like GLMM better capture biological variability and support more precise therapeutic decisions.

Acknowledgement

This research was supported by the Science Fund of the Republic of Serbia, GRANT No. 9393, “Optimization and Prediction in Therapy Treatments of Cancer – OPTIC”, and by the Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, within the framework of Project “Socio-technological aspects of improving the teaching process in the English language in fundamental disciplines”.

References

[1] J. Ivetić, J. Dedeić, S. Milićević, K. Vidojević, and M. Delić, ”Comparative Meta-Analysis of Survival, Risk, and Treatment Efficacy in Immunotherapy for Metastatic Melanoma Using Random-, Fixed-, and Mixed-Effects Models”, *Journal of Clinical Medicine*, 2025.

[2] M. Delić, S. Milićević, K. Vidojević, J. Dedeić, and J. Ivetić, ”Meta-analysis of risks and benefits in immunotherapy treatment of metastatic melanoma”, in *The Book of Abstracts from RAP 2025 Conference*, Heraklion, Crete, Greece, May 2025.

[3] S. Milićević, J. Dedeić, K. Vidojević, J. Ivetić, and M. Delić, ”Model-based meta-analysis of adverse events in metastatic melanoma immunotherapy: The importance of quantifying heterogeneity”, in *Book of Abstracts of RAD 2025 Conference*, Herceg Novi, Montenegro, June 2025.

[4] J. D. Wolchok, V. Chiarion-Sileni, R. Gonzalez, J.J. Grob, P. Rutkowski, et al., ”Long-Term Outcomes With Nivolumab Plus Ipilimumab or Nivolumab Alone Versus Ipilimumab in Patients With Advanced Melanoma”, *Journal of Clinical Oncology*, 2022.

[5] C. Robert, M.S. Carlino, C. McNeil, A. Ribas, J.J. Grob, et al., ”Seven-Year Follow-Up of the Phase III KEYNOTE-006 Study: Pembrolizumab Versus Ipilimumab in Advanced Melanoma”, *Journal of Clinical Oncology*, 2023.

[6] O. Hamid, I. Puzanov, R. Dummer, J. Schachter, A. Daud, et al., ”Final analysis of a randomised trial comparing pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory advanced melanoma”, *European Journal of Cancer*, 2017.

[7] N. Hribernik, M. Boc, J. Ocvirk, J. Knez-Arbeiter, T. Mesti, et al., ”Retrospective analysis of treatment-naive Slovenian patients with metastatic melanoma treated with pembrolizumab—real-world experience.”, *Radiology and Oncology*, 2020.

[8] O. Hamid, C. Robert, A. Daud, F. S. Hodi, W. J. Hwu, et al., ”Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001”, *Annals of Oncology*, 2019.

[9] F. S. Hodi, J. Chesney, A. C. Pavlick, C. Robert, K. F. Grossman, et al., ”Combined nivolumab and ipilimumab versus ipilimumab alone in patients with advanced melanoma: 2-year overall survival outcomes in a multicentre, randomised, controlled, phase 2 trial”, *The Lancet Oncology*, 2016.

[10] N. Hribernik, D. T. Huff, A. Studen, K. Zevnik, Ž. Klaneček, et al., ”Quantitative imaging biomarkers of immune-related adverse events in immune-checkpoint blockade-treated metastatic melanoma patients: a pilot study”, *European Journal of Nuclear Medicine and Molecular Imaging*, 2022.

[11] N. Hribernik, K. Strasek, D. T. Huff, A. Studen, K. Zevnik, et al., ”Role of quantitative imaging biomarkers in an early FDG-PET/CT for detection of immune-related adverse events in melanoma patients: a prospective study”, *Radiology and Oncology*, 2024.

[12] J. Larkin, D. Minor, S. D’Angelo, B. Neyns, M. Smylie, et al., ”Overall Survival in Patients With Advanced Melanoma Who Received Nivolumab Versus Investigator’s Choice Chemotherapy in CheckMate 037: A Randomized, Controlled, Open-Label Phase III Trial”, *Journal of Clinical Oncology*, 2017.

[13] C. Lebbe, N. Meyer, L. Mortier, I. Marquez-Rodas, C. Robert, et al., ”Evaluation of Two Dosing Regimens for Nivolumab in Combination With Ipilimumab in Patients With Advanced Melanoma: Results From the Phase IIIb/IV CheckMate 511 Trial”, *Journal of Clinical Oncology*, 2019.

[14] C. Robert, G. V. Long, B. Brady, C. Dutriaux, A. M. Di Giacomo, et al., "Five-Year Outcomes With Nivolumab in Patients With Wild-Type BRAF Advanced Melanoma", *Journal of Clinical Oncology*, 2020.

[15] H. A. Tawbi, P. A. Forsyth, S. F. Hodi, A. P. Algazi, O. Hamid, et al., ”Long-term outcomes of patients with active melanoma brain metastases treated with combination nivolumab plus ipilimumab (CheckMate 204): Final results of an open-label, multicentre, phase 2 study”, *The Lancet Oncology*, 2021.

[16] J. S. Weber, D. Schadendorf, M. Del Vecchio, J. Larkin, V. Atkinson, et al., ”Adjuvant Therapy of Nivolumab Combined With Ipilimumab Versus Nivolumab Alone in Patients With Resected Stage IIIB-D or Stage IV Melanoma (CheckMate 915)”, *Journal of Clinical Oncology*, 2023.