Network meta-analysis of survival outcomes in the first-line treatment of unresectable or metastatic melanoma: scoping review of methodology

SA72

BIOCAD
Biotechnology Company

A. LAZAREV[1], D. TOLKACHEVA[1], K. SAPOZHNIKOV[1], O. MIRONENKO [1], N. SABLEVA[1]
[1]JSC BIOCAD, Saint Petersburg, Russian Federation

INTRODUCTION

Advanced (unresectable or metastatic) melanoma (aM) has seen dramatic therapeutic transformations over the past decade, with immunotherapy (IT), targeted therapy (TT), and their combinations becoming standard first-line options.
Network meta-analysis (NMA) is increasingly used to synthesize evidence across multiple treatments when direct comparisons are lacking. However, most NMAs in oncology still rely only on published hazard ratios, ignoring the time-dependent nature of treatment effects, which often takes place in comparisons of anti-cancer therapies (e.g., quick responses or diminishing effect on TT versus delayed or durable responses and long-term survivors on IT).
This review systematically examines how NMAs for aM have handled time-to-event outcomes, evaluating methodological choices, their transparency, and the adoption of modern approaches like employing individual patient data (IPD) and flexible survival models.

OBJECTIVE

This study aimed to conduct scoping methodological review of NMAs evaluating the first-line systemic therapies for aM, published between 2016 and 2024, with a focus on how time-to-event outcomes were modelled and whether modern statistical approaches were adopted.
We analysed whether methodological choices reflected evolving standards in survival analysis in oncology or remained anchored in outdated practices.

METHODS

We systematically searched PubMed and Embase up to February 2025 to identify all NMAs comparing three or more first-line systemic therapies for aM, with reported time-to-event outcomes. We sought to evaluate the frequency and transparency of the following key methodological decisions: aggregated versus IPD, fixed (FE) versus random effects (RE), Bayesian versus frequentist framework, proportional versus time-dependent hazards models, standard versus flexible parametric survival distributions, hazard ratios versus other estimands (difference in the milestone survival or restricted mean survival time (RMST)).
We further examined the geographical distribution of lead authors, the impact factors of publishing journals, the temporal evolution of the field, and the consistency of evidence synthesis across studies, with particular attention to variation in inclusion criteria and data extraction methods which could affect methodological rigor and clinical interpretability.
In addition, we assessed the validity of the proportional hazards assumption and evaluated whether authors accounted for differences in mechanisms of action between the compared treatment options (e.g. BRAF/MEK and PD-1/PD-L1 inhibitors)

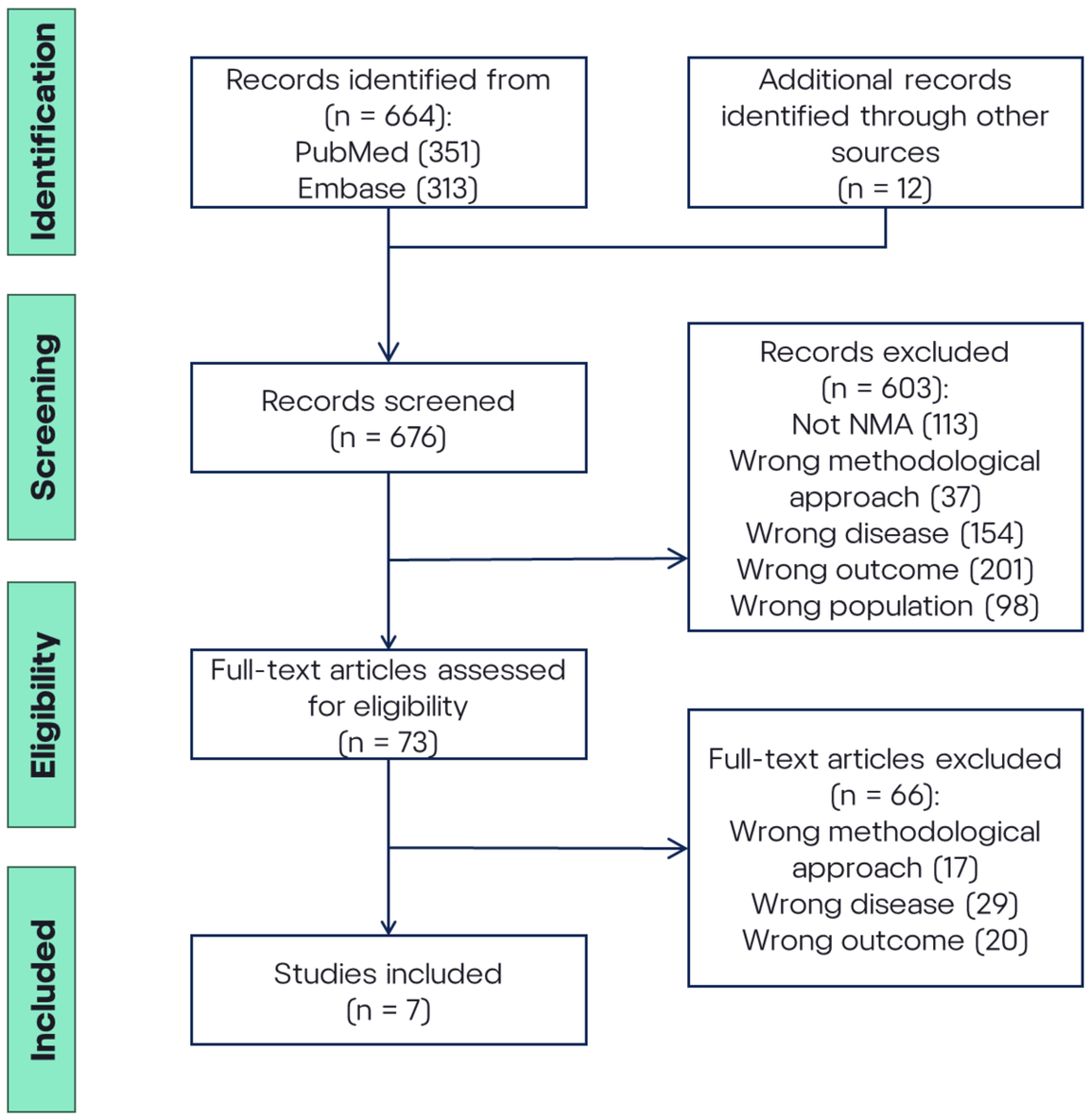
RESULTS

Seven studies were included into current scoping review (fig. 1), each synthesizing data from 7 to 29 randomized controlled trials (RCTs), totaling 45 unique trials and over 40,000 patients.

Among the included NMAs, six relied entirely on published hazard ratios. Only one study (Freeman, 2022) reconstructed pseudo-IPD from Kaplan-Meier curves using Guyot's algorithm and applied six distinct survival models, including Cox PH, difference in the RMST (estimated as the area under the KM curve), piecewise exponential, fractional polynomial, generalized gamma, and the flexible parametric model with Royston-Parmar splines, ultimately selecting the latter for its superior fit and extrapolation capacity. Bayesian methods dominated (six studies), while only Boutros (2023) appeared to use a frequentist approach. FE and RE models were each reported in four studies, with one study (Freeman, 2022) reporting both approaches, yet only Corrie (2022) and Freeman (2022) provided justifications for approach selection. Proportional hazards were formally tested in only three studies (Corrie, Freeman, and Lengyel), using visual inspection of complementary log-log plots or Schoenfeld residuals; the remaining four assumed proportionality without validation. Notably, Amdahl (2016) based its synthesis on two systematic reviews with outdated search cutoffs (2012 and 2015), excluding pivotal trials such as CheckMate 067 and KEYNOTE-006, while Li (2023) included newer agents like relatlimab and high-dose pembrolizumab.

Geographically, five lead authors were from Europe (UK: 2, Italy: 1, Netherlands: 1, Hungary: 1), and per one from the USA and China. The highest-impact journal was Cancer Treatment Reviews (SJR 3.67, Corrie 2022), while the lowest was Oncology and Therapy (SJR 0.898, Amdahl 2016). The number of publications surged after 2022, with four of seven NMAs published in 2022–2024, reflecting growing interest but not improved methodological quality. Crucially, despite the inclusion of immunotherapy combinations with durable, non-proportional effects, only Freeman (2022) modeled time-varying hazard effects, and only one study (Freeman 2022) provided survival curves extending beyond 5 years – highlighting a systemic failure to address the unique temporal dynamics of aM therapies.

Figure 1. *PRISMA flowchart*



CONCLUSIONS

Despite the rapid expansion of NMAs in aM treatment evaluation, the field remains critically underpinned by outdated and opaque methodological practices. The overwhelming reliance on aggregated data, the near-universal absence of proportional hazards testing, and the inconsistent or unjustified selection of FE versus RE models severely limit the validity of survival estimates and the reliability of long-term extrapolation, particularly in an era where immunotherapies induce delayed, non-proportional benefits. The lone exception, Freeman (2022), demonstrated that pseudo-IPD reconstruction combined with flexible parametric modelling (Royston-Parmar splines) enables accurate representation of time-varying treatment effects and supports clinically meaningful long-term projections. Yet this approach remains a rarity, even in high-impact journals. Furthermore, methodological inconsistencies, such as the exclusion of modern therapies, arbitrary aggregation of treatment arms, and inconsistent reporting of BRAF status – introduce bias and hinder reproducibility. Moreover, only a limited number of NMAs have incorporated the most recent first-line therapies, reflecting the time lag between clinical adoption and methodological synthesis. To ensure that NMAs generate evidence fit for clinical decision-making and health policy, future studies must adopt pseudo-IPD synthesis, rigorously test proportional hazards assumptions, justify model selection transparently, and prioritize flexible parametric survival models capable of capturing the complex temporal dynamics of modern melanoma therapies. Only then network meta-analysis can evolve from a descriptive tool into a truly predictive and actionable instrument for healthcare decision-making.

CONTACT INFORMATION

Andrei Lazarev
JSC BIOCAD
E-mail: lazarevaa@biocad.ru

REFERENCES