European Journal of Surgical Oncology 52 (2026) 111915

Contents lists available at ScienceDirect

European Journal of Surgical Oncology

journal homepage: www.ejso.com

Comparison of the effectiveness of neoadjuvant chemotherapy versus upfront surgery for osteosarcoma: A target trial emulation study

Guodong Zhong [a,b,1], Honglin Gu [c,1], Zijun Huo [d,1], Mingxi Zhu [a,b], Julian A. Gingold [e], Dung-Fang Lee [f,g,h], Gang Huang [a,b], Jingnan Shen [a,b,*,2], Jian Tu [a,b,**,2]

[a] Department of Musculoskeletal Oncology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China
[b] Guangdong Provincial Key Laboratory of Orthopedics and Traumatology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China
[c] Department of Spine Surgery, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, No. 106 Zhongshan Road II, Guangzhou, Guangdong Province, 510080, China
[d] Department of Endocrinology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China
[e] Department of Obstetrics & Gynecology and Women's Health, Einstein/Montefiore Medical Center, Bronx, NY, 10461, United States
[f] Department of Integrative Biology and Pharmacology, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, 77030, United States
[g] The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at Houston, Houston, TX, 77030, United States
[h] Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, 77030, United States

ARTICLE INFO

Keywords:
Osteosarcoma
Therapeutics
Neoadjuvant therapy
Surgery timing
Survival

ABSTRACT

Introduction: Most studies recommend a treatment sequence involving neoadjuvant chemotherapy followed by surgery and adjuvant chemotherapy (NAC first) for osteosarcoma, yet direct comparisons with upfront surgery (surgery first) are lacking, and previous comparative analyses revealed no significant difference in overall survival (OS) between the two strategies.

Materials and methods: Using the target trial emulation (TTE) framework and the Surveillance, Epidemiology, and End Results (SEER) database, we compared NAC first versus surgery first in patients with osteosarcoma. Propensity score matching (PSM) and inverse probability of treatment weighting (IPTW) were employed to control for baseline confounding. The primary analysis aimed to estimate the intention-to-treat (ITT) effect to evaluate the comparative effectiveness of which initial treatment strategy is associated with a better prognosis.

Results: Among 831 eligible patients, 152 were assigned to the surgery-first group, and 679 were assigned to the NAC-first group. After PSM (121 pairs), the NAC-first group had significantly higher 5-year OS (77.7% [95% CI: 70.6%–85.5%] versus 61.3% [95% CI: 53.0%–70.8%], $p = 0.006$) and an increased risk of death in the surgery-first group (HR = 1.907, 95% CI: 1.172–3.103; $p = 0.009$). ITT and per-protocol (PP) analyses consistently corroborated these findings. IPTW analysis of the entire cohort showed a non-significant trend toward improved survival with NAC first (HR = 1.084, 95% CI: 0.992–1.186; $p = 0.076$).

Conclusion: This study is the first to employ the TTE framework to compare treatment sequences for osteosarcoma. Our findings suggest that initiating treatment with NAC is associated with increased OS rates compared with initiating treatment with upfront surgery, providing empirical evidence that aligns with current recommendation in the National Comprehensive Cancer Network (NCCN) guidelines.

* Corresponding author. The First Affiliated Hospital of Sun Yat-sen University, No.58 Zhongshan II Road, Guangzhou, Guangdong Province, PR China. Guangdong Provincial Key Laboratory of Orthopedics and Traumatology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China.
** Corresponding author. Department of Musculoskeletal Oncology, The First Affiliated Hospital of Sun Yat-sen University, No.58 Zhongshan II Road, Guangzhou, Guangdong Province, China.
E-mail addresses: zhonggd@mail2.sysu.edu.cn (G. Zhong), guhonglin@gdph.org.cn (H. Gu), huozj@mail.sysu.edu.cn (Z. Huo), zhumx9@mail2.sysu.edu.cn (M. Zhu), Jgingold@gmail.com (J.A. Gingold), dung-fang.lee@uth.tmc.edu (D.-F. Lee), hgang@mail.sysu.edu.cn (G. Huang), shenjn@mail.sysu.edu.cn (J. Shen), tujian3@mail.sysu.edu.cn (J. Tu).
[1] Guodong Zhong, Honglin Gu and Zijun Huo contributed equally to this work.
[2] Jingnan Shen and Jian Tu contributed equally as supervising authors.

https://doi.org/10.1016/j.ejso.2026.111915
Received 13 March 2026; Received in revised form 19 May 2026; Accepted 29 May 2026
Available online 2 June 2026

1. Introduction

Osteosarcoma, which is the most common primary malignant bone tumor, is highly aggressive and prone to pulmonary metastasis [1]. Since the introduction of chemotherapy in the 1970s [2], the overall survival (OS) rate for patients with localized osteosarcoma has increased from 10–20% to 60–80% [3–6]. Currently, the guideline-recommended first-line treatment for osteosarcoma remains neoadjuvant chemotherapy (NAC) followed by surgery and adjuvant chemotherapy [7]. However, the literature underlying these guideline-based recommendations primarily consists of studies demonstrating that different preoperative NAC regimens and cycles can achieve relatively stable survival rates [8–13]. Yet, direct comparisons between a strategy of initial NAC followed by surgery (NAC first) and a strategy of immediate surgery (surgery first), both followed by adjuvant chemotherapy, are lacking.

Although direct comparisons between the two strategies exist, including one randomized controlled trial (RCT) and two retrospective studies, which generally reported no significant difference in OS, the evidence remains inconclusive. The pivotal RCT [14], while methodologically superior, was critically undermined by slow patient accrual and enrolled only 100 total patients but identified largely comparable results. As explicitly stated in the trial report, the original design of the study was amended and it ultimately lacked statistical power (only 63% power to detect a 15% inferior outcome) to reliably exclude a clinically significant difference in survival, thus limiting identification of clinically meaningful differences associated with either strategy. The largest retrospective cohort study [15], which identified over 700 patients, suggested the potential for improved OS with a surgery-first management approach. Retrospective analyses [15,16] are inherently limited by methodological concerns, including immortal time bias, selection bias, and residual confounding due to unmeasured factors, thereby precluding a causal inference [17–19]. Consequently, the optimal sequence of therapy remains unresolved and inadequately addressed in the literature.

By explicitly designing studies to emulate a hypothetical randomized trial, target trial emulation (TTE) [20–22] offers a rigorous framework for causal inference from observational data. Its key advantages include preventing avoidable biases such as immortal time and selection bias through the alignment of eligibility, treatment assignment, and follow-up at time zero; addressing causal questions around actionable interventions; guiding appropriate confounder adjustment and analysis planning; and increasing interpretability through structured, protocol-driven emulation that increases the credibility and clinical relevance of observational findings.

As a method to address this evidence gap, we employed the TTE framework and the Surveillance, Epidemiology, and End Results (SEER) database to simulate an RCT, aiming to provide more robust and adequately powered evidence of comparative effectiveness.

2. Materials and methods

2.1. Data source

We obtained data from the SEER database [23]. Patient data were extracted using SEER*Stat software (version 9.0.41). The SEER 17 registries (2000–2022), which cover approximately 26.5% of the U.S. population (based on the 2020 census), provide a broadly representative sample for cancer research. Follow-up data were available through the end of 2022.

2.2. Eligibility criteria

The inclusion criteria included (1) a pathologically confirmed diagnosis of extremity osteosarcoma between 2007 and 2017 and (2) the initiation of either surgery or NAC within 90 days after diagnosis [15]. The year 2007 was selected as the starting point because comprehensive records regarding the sequence of treatment (NAC first versus surgery first) became available from that time onward. Furthermore, the use of this time frame ensured that all enrolled patients could have at least 5 years of follow-up. Patients were excluded if they had incomplete data on tumor size or the time from the diagnosis to initial treatment or if they had concurrent or prior malignancies.

The variables collected in this study included age, sex, race and origin, year of diagnosis, tumor location, tumor size, histology, tumor stage, and surgical approach.

2.3. Assignment procedures

All eligible patients were assigned to one of the two treatment strategies: those in the NAC-first group initiated NAC within 90 days after diagnosis, followed by surgery and adjuvant chemotherapy; those in the surgery-first group underwent surgery within 90 days after diagnosis, followed by adjuvant chemotherapy. All patients were followed from the time of enrollment (time zero) until death, loss to follow-up, or the end of the fifth year, whichever occurred first (Fig. 1). Critically, time zero was strictly defined as the date of diagnosis for all patients, regardless of which treatment strategy they ultimately received. This alignment of the survival clock at diagnosis—rather than at treatment initiation—is a foundational element of the TTE framework and is essential for preventing immortal time bias [20]. To ensure uniformity in recording and analysis, a standardized time unit was applied, with each month defined as 30 days and each year defined as 360 days.

2.4. Outcome

The primary outcome was 5-year OS, defined as the time from diagnosis until death from any cause, loss to follow-up, or until the end of the follow-up period, whichever occurred first.

2.5. Statistical analysis

For continuous variables, the normality of distribution was assessed using the Shapiro–Wilk test. Variables conforming to a normal distribution are presented as the means ± standard deviations (SDs), while those not conforming to a normal distribution are described as medians and interquartile ranges (IQRs). Categorical variables are summarized as frequencies and percentages.

This study commenced with an unadjusted preliminary analysis of the collected data, which included an OS analysis and multivariable Cox regression analysis. The primary analysis was conducted based on the intention-to-treat (ITT) principle to compare the effects of the treatment strategies assigned at baseline, regardless of whether individuals subsequently adhered to their assigned strategy.

Two distinct analytical methods were employed to address potential confounding and ensure comparability of the baseline characteristics between groups: propensity score matching (PSM) [24] and inverse probability of treatment weighting (IPTW) [25]. For PSM, nearest neighbor matching was performed at a 1:1 ratio with a caliper width of 0.2 standard deviations of the propensity score logit. For IPTW, the inverse probability of treatment weights was derived from propensity scores and stabilized by truncation at the 1st and 99th percentiles. The PSM-based ITT analysis of the matched cohort was pre-specified as the primary analysis for this study, as it provides an estimate of the average treatment effect on the treated (ATT) with optimal confounding control. The IPTW-based ITT analysis of the full cohort was pre-specified as a confirmatory analysis to assess the robustness of findings under an alternative weighting strategy and to utilize the full sample size. Balance was assessed using standardized mean differences (SMDs) and visualized in Love plots for both the PSM-adjusted cohort and the IPTW-adjusted cohort, with SMD values less than 0.1 indicating adequate balance. Both approaches were subsequently incorporated into a clone-censor-weighting (CCW) framework [20] to avoid immortal time

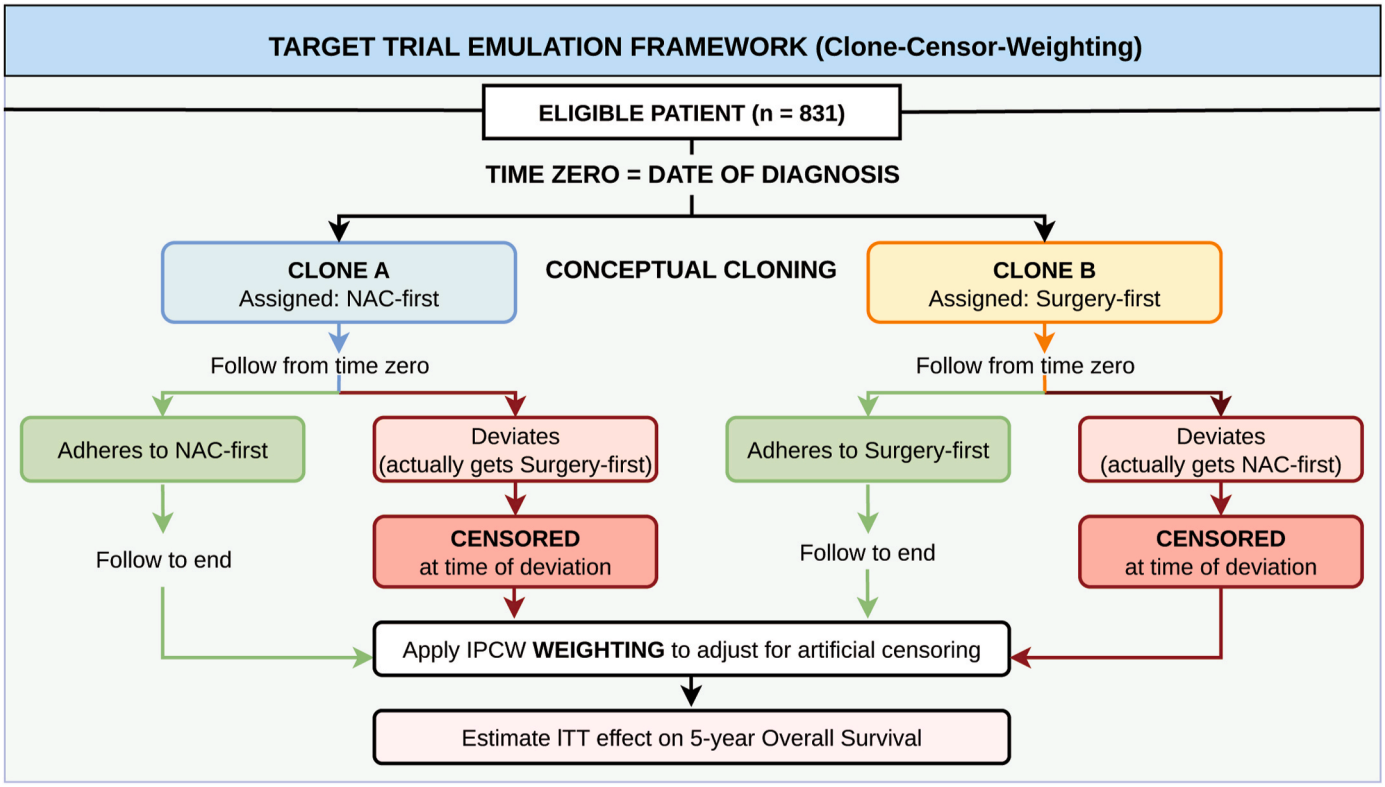

Fig. 1. Schematic illustration of the target trial emulation (TTE) framework using clone-censor-weighting (CCW). At time zero (date of diagnosis), each eligible patient is conceptually cloned into two identical copies. One clone is assigned to the NAC-first strategy, and the other to the Surgery-first strategy. Both clones are followed forward in time from time zero. If the actual patient's treatment trajectory deviates from a clone's assigned strategy, that clone is artificially censored at the time of deviation. Inverse probability of censoring weighting (IPCW) is applied to adjust for potential selection bias introduced by this artificial censoring. The overall survival (OS) is assessed at 5 years from time zero. This framework ensures that immortal time bias is avoided, as the survival clock starts at diagnosis for both strategies, and deaths occurring during the pretreatment window are appropriately attributed to both arms in the intention-to-treat (ITT) analysis.

bias. In this step, the inverse probability of censoring weights (IPCW) [26], which was truncated at the 1st and 99th percentiles, was applied to further adjust for informative censoring. Following these adjustments, Kaplan–Meier (KM) curves were constructed from the weighted populations to visualize survival distributions, and weighted Cox proportional hazards models were established via data fitting for statistical inference. However, as standard log-rank p-values are not routinely reported for weighted survival curves because of methodological limitations, formal hypothesis testing was performed using weighted Cox proportional hazards models rather than through a direct comparison of the weighted Kaplan–Meier estimates.

In addition, a per-protocol (PP) effect analysis was performed to evaluate the biological effect under ideal conditions by including only those patients who strictly adhered to the assigned treatment strategy throughout the study period. The PP analysis was defined from the start of treatment rather than the time of enrollment. This analysis provides insights into the efficacy of each strategy when implemented as originally intended while also reducing bias related to treatment delays or nonadherence.

Finally, subgroup analyses were performed on the basis of the ITT population to assess the consistency of the treatment effect between NAC first and surgery first across various patient subsets. These analyses were pre-specified as exploratory and were intended to generate hypotheses regarding potential effect modification. Subgroups were defined based on the following covariates: age, sex, race and ethnicity, year of diagnosis, primary tumor site, tumor size, histology, disease stage, and surgical approach. Specifically, age was categorized as 0–10 years, 10–20 years, and >20 years, and tumor size was classified into five groups: 0–50 mm, 51–100 mm, 101–150 mm, 151–200 mm, and >200 mm.

To assess the potential impact of unmeasured confounding on the primary results, we calculated the E-value for the observed hazard ratios [27]. The E-value quantifies the minimum strength of association that an unmeasured confounder would need to have with both the treatment assignment and the outcome to fully explain away the observed association, conditional on the measured covariates.

All the statistical analyses and data visualization were performed using R software (version 4.4.0), and the final figures were refined with Adobe Illustrator 2024 for enhanced clarity and presentation. A p value of less than 0.05 was considered to indicate statistical significance.

3. Results

3.1. Patient characteristics

A total of 831 patients were included in this study, with 152 in the surgery-first group and 679 in the NAC-first group (Table 1). The baseline characteristics of the two groups were imbalanced. Patients in the surgery-first group were older (median age: 32 years vs. 14 years, $p = 0.000$) and had smaller tumors (median size: 91 mm vs. 99 mm, $p = 0.009$); moreover, differences were observed in tumor histology ($p = 0.000$) and the surgical approaches employed ($p = 0.001$). The median time from diagnosis to treatment initiation was 11.5 days (IQR: 0–26.25) in the surgery-first group and 13 days (IQR: 8–22) in the NAC-first group. All patients initiated their assigned treatment strategy within the 90-day eligibility window (Table S1).

Table 1
Baseline characteristics of the patients enrolled in this study.

Variable	Total (n = 831)	Surgery First (n = 152)	NAC First (n = 679)	P value
Age (median, IQR)	15 (12–20)	32 (16.75–47.25)	14 (11–18)	0.000
Sex (n, %)				0.401
Male	497 (59.8%)	96 (63.2%)	401 (59.1%)	
Female	334 (40.2%)	56 (36.8%)	278 (40.9%)	
Race and origin (n, %)				0.940
Hispanic	263 (28.4%)	45 (29.6%)	218 (32.1%)	
NHW	359 (43.2%)	68 (44.7%)	291 (42.9%)	
NHB	110 (13.2%)	20 (13.2%)	90 (13.3%)	
Other	99 (11.9%)	19 (12.5%)	80 (11.8%)	
Year of diagnosis (n, %)				0.126
2007–2009	209 (25.2%)	49 (32.2%)	160 (23.6%)	
2010–2012	212 (25.5%)	33 (21.7%)	179 (26.4%)	
2013–2015	247 (29.7%)	45 (29.6%)	202 (29.7%)	
2016–2017	163 (19.6%)	25 (16.4%)	138 (20.3%)	
Location (n, %)				0.789
Lower limb	699 (84.1%)	133 (87.5%)	566 (83.4%)	
Upper limb	132 (15.9%)	29 (19.1%)	113 (16.6%)	
Size (median, IQR)	97 (72–130)	91 (60.0–120.5)	99 (75–133)	0.009
Histology (n, %)				0.000
Conventional	559 (67.3%)	83 (54.6%)	476 (70.1%)	
Chondroblastic	121 (14.6%)	22 (14.5%)	99 (14.6%)	
Other	151 (18.2%)	47 (30.9%)	104 (15.3%)	
Stage [a] **(n, %)**				0.603
I–II	625 (75.2%)	119 (78.3%)	506 (74.5%)	
III–IV	165 (19.9%)	27 (17.8%)	138 (20.3%)	
Unknown	41 (4.9%)	6 (3.9%)	35 (5.2%)	
Surgery (n, %)				0.001
Limb salvage	617 (74.2%)	110 (72.4%)	507 (74.7%)	
Amputation	207 (24.9%)	37 (24.3%)	170 (25.0%)	
Unknown	7 (0.8%)	5 (3.3%)	2 (0.3%)	

NAC, neoadjuvant chemotherapy; IQR, interquartile range; NHW, non-Hispanic white; NHB, non-Hispanic black.
[a] American Joint Committee on Cancer (AJCC) Staging System.

3.2. Assessment of the treatment effect via the PSM method

3.2.1. Construction of the PSM cohort

We first employed PSM to balance the intergroup differences. After 1:1 matching, a total of 121 matched pairs were retained. The baseline characteristics of the matched cohort were generally well balanced, with all p values exceeding 0.05 (Table 2). Balance diagnostics are presented in the Love plot (Fig. S1A). Before matching, substantial imbalances were observed. After 1:1 propensity score matching, the SMDs for most covariates were reduced to below the conventional threshold of 0.1, while a small number slightly exceeded this threshold, which may be attributed to the substantial baseline imbalance prior to matching (Table S2). These results indicate that the baseline characteristics were adequately balanced and that the groups were comparable for subsequent comparative analyses.

Table 2
Baseline characteristics of patients enrolled in the study (PSM 1:1).

Variable	Total (n = 242)	Surgery First (n = 121)	NAC First (n = 121)	P value
Age (median, IQR)	21 (15–40.75)	26 (15–38)	18 (15–42)	0.370
Sex (n, %)				0.599
Male	147 (60.7%)	76 (62.8%)	71 (58.7%)	
Female	95 (39.3%)	45 (37.2%)	50 (41.3%)	
Race and origin (n, %)				0.915
Hispanic	76 (31.4%)	38 (31.4%)	38 (31.4%)	
NHW	104 (43.0%)	54 (44.6%)	50 (41.3%)	
NHB	27 (11.2%)	12 (9.9%)	15 (12.4%)	
Other	35 (14.5%)	17 (14.0%)	18 (14.9%)	
Year of diagnosis (n, %)				0.809
2007–2009	79 (32.6%)	36 (29.8%)	43 (35.5%)	
2010–2012	53 (21.9%)	27 (22.3%)	26 (21.5%)	
2013–2015	70 (28.9%)	37 (30.6%)	33 (27.3%)	
2016–2017	40 (16.5%)	21 (17.4%)	19 (15.7%)	
Location (n, %)				0.853
Lower limb	208 (86.0%)	105 (86.8%)	103 (85.1%)	
Upper limb	34 (14.0%)	16 (13.2%)	18 (14.9%)	
Size (median, IQR)	90 (67.25–120)	96 (67–120)	87 (68–115)	0.621
Histology (n, %)				0.725
Conventional	136 (56.2%)	69 (57.0%)	67 (55.4%)	
Chondroblastic	33 (13.6%)	18 (14.9%)	15 (12.4%)	
Other	73 (30.2%)	34 (28.1%)	39 (32.2%)	
Stage (n, %)				0.477
I–II	201 (83.1%)	97 (80.2%)	104 (86.0%)	
III–IV	32 (13.2%)	19 (15.7%)	13 (10.7%)	
Unknown	9 (3.7%)	5 (4.1%)	4 (3.3%)	
Surgery (n, %)				0.642
Limb salvage	186 (76.9%)	90 (74.4%)	96 (79.3%)	
Amputation	52 (21.5%)	29 (24.0%)	23 (19.0%)	
Unknown	4 (1.7%)	2 (1.7%)	2 (1.7%)	

NAC, neoadjuvant chemotherapy; IQR, interquartile range; NHW, non-Hispanic white; NHB, non-Hispanic black.

3.2.2. Initial analysis of the matched cohort

The survival analysis revealed that the 5-year OS rates were 77.7% (95% confidence interval [CI]: 70.6%–85.5%) in the NAC-first group and 61.3% (95% CI: 53.0%–70.8%) in the surgery-first group, and the difference was statistically significant (p = 0.006) (Fig. 2A). According to the Cox proportional hazards model, the surgery-first group exhibited a significantly higher risk of death than the NAC-first group (hazard ratio [HR] = 1.907, 95% CI: 1.172–3.103; p = 0.009), indicating a 90.7% increase in the mortality risk for patients who received surgery first (Table S3, Fig. 2B).

3.2.3. ITT analysis

In the ITT analysis, the 5-year OS rates were 71.7% (95% CI: 66.1%–77.7%) in the NAC-first group and 68.4% (95% CI: 62.6%–74.8%) in the surgery-first group (Fig. 2C). Compared with the NAC-first group, the surgery-first group had a significantly higher risk of death (HR = 1.119, 95% CI: 1.021–1.226; p = 0.017), corresponding to an 11.9% increase in the mortality risk (Table S3, Fig. 2D).

For the primary ITT analysis (HR = 1.119), the calculated E-value was 1.48 (lower 95% CI bound E-value: 1.16). This suggests that to fully explain away the observed association, an unmeasured confounder would need to increase the likelihood of upfront surgery and the risk of death by approximately 50%, conditional on measured covariates.

3.2.4. PP sensitivity analysis

The KM curve (Fig. 2E) showed that the per-protocol 5-year OS rates were 77.1% (95% CI: 69.9%–85.1%) in the NAC-first group and 64.2% (95% CI: 54.8%–75.1%) in the surgery-first group, and the difference

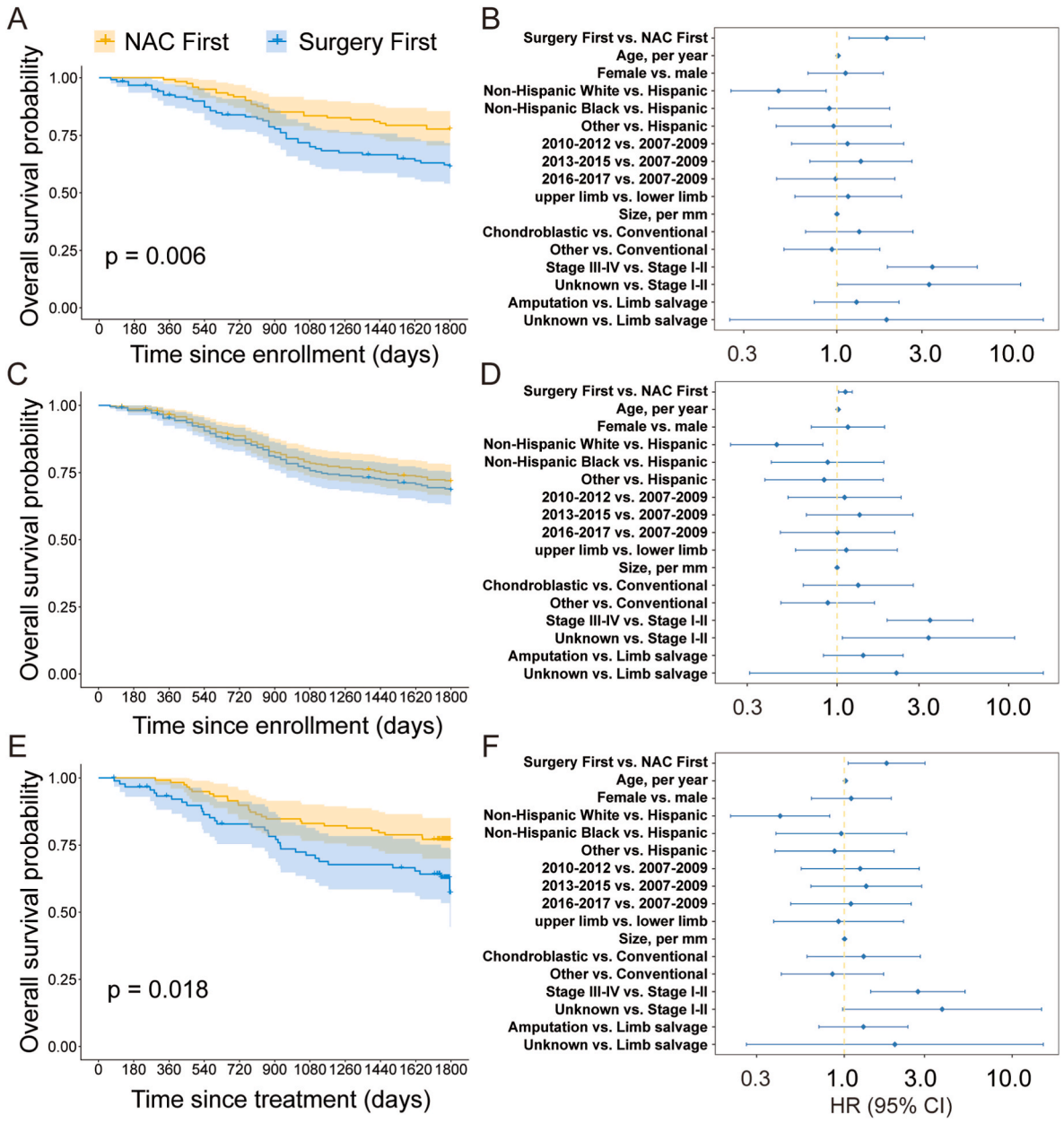

Fig. 2. Assessment of the treatment effect using propensity score matching (PSM) to address potential confounding bias. Analyses included an initial analysis of the matched cohort (A, B), intention-to-treat (ITT) analysis (C, D), and per-protocol (PP) analysis as a sensitivity analysis (E, F).

between the groups was statistically significant (p = 0.018). Cox regression analysis revealed a significantly increased risk of death in the surgery-first group compared with the NAC-first group (HR = 1.785, 95% CI: 1.054–3.022; p = 0.031) (Table S3, Fig. 2F).

3.2.5. Subgroup analysis

In the overall study population, the surgery-first regimen was associated with a significantly increased risk of death compared with the NAC-first regimen. We performed a total of nine subgroup analyses to identify potential modifiers of the effect. These exploratory analyses revealed a nominal interaction by sex (p for interaction = 0.034), with the point estimate for the HR suggesting a more pronounced increase in risk observed among male patients but not among female patients (Fig. S2). However, given the multiple subgroup comparisons performed and the absence of multiplicity adjustment, this finding may represent a Type I error and should be interpreted with extreme caution. No significant interactions were observed for any other baseline characteristics examined. The consistency of the overall treatment effect across most subgroups supports the robustness of the primary findings, but the observed sex-specific difference requires validation in independent cohorts before any clinical inferences can be drawn. For all other subgroups, although the 95% CIs for the HRs crossed 1 in some analyses, indicating nonsignificant differences in OS within those specific subgroups, no significant interactions were found (Fig. S2). This pattern implies that the treatment effect is largely consistent across these subgroups and that the observed variations are more likely attributable to limited statistical power or random chance rather than to true heterogeneous treatment effects.

3.3. Confirmatory analysis using IPTW

3.3.1. Unadjusted preliminary analysis

We initially performed an analysis using the unadjusted data, which included the KM and multivariable Cox regression analyses. The KM

curve showed a 5-year OS rate of 68.2% (95% CI: 64.7%–71.8%) for the NAC-first group, whereas it was 59.7% (95% CI: 52.3%–68.2%) for the surgery-first group (p = 0.032) (Fig. S3A). The Cox proportional hazards model results indicated an HR of 1.359 (95% CI: 0.969–1.907) for the surgery-first group compared with the NAC-first group, with a p value of 0.076 (Table S4, Fig. S3B).

3.3.2. ITT analysis

We applied IPTW to adjust for intergroup imbalances and used CCW to account for immortal time bias. The Love plot (Fig. S1B) showed an adequate balance of baseline characteristics after weighting, with an SMD less than 0.1 for all covariates (Table S2). After adjustment, the 5-year OS rate was marginally higher in the NAC-first group (66.1%, 95% CI: 61.4%–71.2%) than in the surgery-first group (64.6%, 95% CI: 58.4%–71.6%) (Fig. S3C). The adjusted multivariable Cox regression results suggested a trend toward an increased risk of death in the surgery-first group compared with the NAC-first group (HR = 1.084, 95% CI: 0.992–1.186; p = 0.076) (Table S4, Fig. S3D).

3.3.3. PP sensitivity analysis

A PP analysis was performed to assess the treatment effect under ideal conditions. The KM curve showed a lower 5-year OS rate in the surgery-first group (65.2%, 95% CI: 54.0%–78.7%) than in the NAC-first group (69.3%, 95% CI: 65.5%–73.2%) (Fig. S3E). Similarly, compared with the NAC-first group, the surgery-first group exhibited a numerically greater, but not statistically significant, risk of death (HR = 1.367, 95% CI: 0.878–2.128; p = 0.166), corresponding to a 36.7% increase in mortality risk (Table S4, Fig. S3F).

3.3.4. Subgroup analysis

In the overall study population, all three analytical approaches (unadjusted, ITT, and PP) yielded HR point estimates greater than 1, consistently indicating a trend toward an increased risk of death in the surgery-first group. Although the 95% CIs included the null value (HR = 1) in each case, suggesting nonsignificant trends, the primary ITT analysis revealed a borderline significant result (p = 0.076). In the same nine subgroup analyses, none of the interactions between treatment strategies and baseline characteristics were statistically significant (all p values for interaction >0.05), indicating that the observed trend of an increased risk associated with the surgery-first strategy was generally consistent across all the subgroups (Fig. S4).

3.4. Systematic literature review and integration of the evidence

Currently, for patients with high-grade osteosarcoma, the NCCN guidelines still recommend a treatment strategy consisting of NAC followed by wide excision and subsequent adjuvant chemotherapy as the optimal approach ("Category 1 On the basis of high-level evidence, there is uniform NCCN consensus that the intervention is appropriate") [7].

However, none of the studies cited by the NCCN guidelines directly compare the difference in survival outcomes between these two treatment strategies.

In contrast, all three existing studies in which survival was directly compared between the two strategies did not demonstrate superior overall survival in patients who received NAC than in those who underwent immediate surgery (Table 3). In an RCT [14], the 5-year OS was 79% in the surgery-first group and 76% in the NAC-first group (p = 0.6). A study by Xu et al. [16]. Focusing on nonmetastatic high-grade pelvic osteosarcoma also revealed that patients who received delayed surgery after NAC did not experience a survival benefit compared with those who received immediate surgery followed by adjuvant chemotherapy. The 5-year OS was 43% in the NAC-first group and 40% in the surgery-first group (p = 0.709). Another retrospective study utilizing the SEER database analyzed 5-year OS in patients aged 5 to 29 years with extremity osteosarcoma [15]. Surprisingly, the results indicated a survival advantage in the surgery-first group compared with the NAC-first group (74.1% vs. 67.2%).

Table 3
Comparison of 5-year OS between patients with osteosarcoma who received NAC first and those who received surgery first.

Research type	NAC First		Surgery First		P value	References
	N	5-year OS	N	5-year OS		
RCT	45	76%	55	79%	0.6	Goorin et al. (2003)
Retrospective study	56	43%	33	40%	0.709	Xu et al. (2019)
Retrospective study	725	67.2%	67	74.1%	NA	Danese et al. (2025)

OS, overall survival; NAC, neoadjuvant chemotherapy; N, number of patients; RCT, randomized controlled trial; NA, not applicable.

Collectively, these findings suggest that for both patients with pelvic and extremity osteosarcoma, the NAC-first strategy was not associated with a better prognosis compared with the surgery-first approach, which is inconsistent with the recommendations in the current NCCN guidelines.

4. Discussion

The NCCN osteosarcoma guidelines currently recommend a strategy of chemotherapy followed by surgery. For this reason, an RCT comparing the outcomes of osteosarcoma patients who undergo surgery first versus NAC is particularly ethically and logistically challenging to complete, in addition to being particularly costly and often slow. [28–30]. For certain rare diseases, participant recruitment also presents considerable challenges and could even lead to early termination of the study [31].

The TTE study approach provides a robust framework for estimating causal effects from observational data by explicitly designing studies to replicate a hypothetical RCT [32] and overcome the ethical, financial and logistical constraints encountered in true RCTs. Unlike conventional retrospective studies [33], which are frequently susceptible to avoidable methodological flaws, the TTE framework mandates the prespecification of key trial components such as eligibility criteria, treatment strategies, and a clear time zero. This rigorous structure proactively addresses and mitigates common sources of bias, such as immortal time bias and selection bias. By aligning the observational analysis with the principles of randomized study design, TTE significantly increases the quality and credibility of causal inference from nonexperimental data [34].

This study used a large observational database (the SEER 17 registries) within a TTE framework to address the question of whether a difference in 5-year OS of osteosarcoma patients exists between surgery-first and NAC-first strategies. After applying PSM to control for confounding variables, the ITT analysis based on the matched cohort showed a significantly increased risk of death in the surgery-first group (p = 0.017). The sensitivity analyses yielded consistent results. Subgroup analyses revealed that the surgery-first strategy was associated with significantly worse survival across all subgroups except among female patients, for whom no significant difference was observed.

Additionally, IPTW was used to balance baseline characteristics between groups. The primary ITT analysis indicated a higher risk of death in the surgery-first group than in the NAC-first group. Although this difference did not reach statistical significance, a consistent trend was observed. These findings were further supported by the results of multiple subgroup analyses.

In summary, we attempted to emulate a target trial using observational data, adhering as closely as possible to its key design component [35,36]. Therefore, these findings represent a methodological improvement over conventional retrospective studies, although causal inference remains limited [18,37,38]. The results provide supportive evidence supporting the current NCCN guidelines recommending NAC

prior to surgery.

Our observation of a significant survival advantage with NAC-first in the PSM overlap cohort contrasts with the null findings of the only randomized controlled trial to date (POG-8651) [14] and a recent comprehensive meta-analysis [39] of 25 studies involving 4867 patients. In POG-8651, the 5-year event-free survival was 69% for immediate surgery versus 61% for presurgical chemotherapy ($p = 0.8$), leading the authors to conclude that there was no advantage for presurgical chemotherapy. Similarly, the meta-analysis reported a pooled risk ratio for OS of 1.09 (95% CI: 0.95–1.24; $p = 0.225$), indicating no statistically significant difference between strategies.

Several factors may explain the discrepancy between these prior reports and our findings. First, the POG-8651 trial, despite its methodological rigor as an RCT, was critically underpowered. As explicitly stated in the original trial report, the study had only 63% power to detect a 15% absolute difference in outcome, meaning that a clinically meaningful benefit of NAC could not be reliably excluded. Second, the meta-analysis pooled studies with substantial heterogeneity in treatment eras, chemotherapeutic protocols, and analytical methods. The majority of included observational studies did not employ methods to address immortal time bias, which is inherent to comparisons of sequences where one strategy involves a delay to definitive surgery. Our application of the TTE framework with CCW specifically mitigates this bias, potentially revealing a signal that was attenuated in prior analyses.

The clinical rationale for NAC-first is well established: early systemic therapy may eradicate micrometastases, the preoperative interval permits surgical planning and prosthesis customization, tumor shrinkage can facilitate complete resection with clearer margins, and the degree of chemotherapy-induced necrosis provides prognostic information that can guide adjuvant therapy decisions [3,40–43].

However, this study exhibits several limitations. First, the SEER database lacks data on specific chemotherapy regimens and cycles. Although regimens have remained largely stable over the study period [1,5,7,44–49], we cannot exclude confounding by differences in chemotherapy intensity or composition, which constitutes a source of potential residual confounding. Second, SEER does not capture local recurrence or distant metastasis, limiting the evaluation of event-free survival; therefore, we intentionally focused on overall survival as a reliable, patient-centric endpoint.

The primary ITT HR of 1.119 represents a modest effect (~12% relative mortality increase), which may still be clinically meaningful given the survival plateau in osteosarcoma [1,6], where even small gains (improved 6-year overall survival by 8 absolute percentage points) have been considered practice-changing [46]. Notably, this estimate is substantially smaller than the crude HR of 1.907 from the PSM-matched cohort without full CCW adjustment, illustrating the attenuation achieved by the TTE framework. We therefore caution that the effect sizes should not be interpreted as evidence of a large treatment effect but rather as observational support for the existing NCCN guideline.

Confounding by indication remains an inherent limitation, as SEER lacks data on performance status, comorbidities, and clinical reasoning behind treatment selection. For example, patients with contraindications to methotrexate may be preferentially triaged to upfront surgery and have a worse prognosis. The E-value of 1.48 underscores this vulnerability, indicating that an unmeasured confounder increasing both the odds of upfront surgery and the hazard of death by approximately 50% could nullify the point estimate.

Given the observational design, our estimates should be interpreted as the association under the assumptions of no unmeasured confounding and correct model specification, rather than definitive causal effects. Nevertheless, several considerations support the credibility of our findings. First, the IPTW analysis of the full cohort yielded a consistent trend (HR = 1.084). Second, the PP analysis suggests that among those who actually completed assigned strategies, the benefit was more pronounced. Third, biological plausibility favors NAC: early eradication of micrometastases and improved surgical margins are well-established mechanisms. Given the low probability of future RCTs, these observational findings represent the best available evidence to guide clinical decision-making, albeit with the explicit caveat that unmeasured confounding cannot be fully excluded.

5. Conclusion

Our analysis showed that initiating treatment with NAC was associated with significantly increased OS rates compared with initiating treatment with upfront surgery after statistical adjustment within a TTE design. By mitigating biases inherent in prior observational designs, these findings provide more robust evidence favoring the current NCCN-recommended strategy of NAC-first. While residual confounding cannot be fully excluded in the absence of randomization, this analysis suggests a strong association between initiating NAC and improved long-term survival.

Contribution author(s)

Study concepts: Guodong Zhong; Honglin Gu; Zijun Huo.

Study design: Guodong Zhong; Honglin Gu; Zijun Huo; Gang Huang; Jingnan Shen; Jian Tu.

Data acquisition: Guodong Zhong; Honglin Gu; Zijun Huo.

Quality control of data and algorithms: Guodong Zhong; Honglin Gu; Zijun Huo.

Data analysis and interpretation: Guodong Zhong; Honglin Gu; Zijun Huo; Mingxi Zhu.

Statistical analysis: Guodong Zhong; Honglin Gu; Zijun Huo; Mingxi Zhu.

Manuscript preparation: Guodong Zhong; Honglin Gu; Zijun Huo; Mingxi Zhu.

Manuscript editing: Guodong Zhong; Honglin Gu; Zijun Huo; Julian A. Gingold; Dung-Fang Lee; Gang Huang; Jingnan Shen; Jian Tu.

Manuscript review: Guodong Zhong; Honglin Gu; Zijun Huo; Julian A. Gingold; Dung-Fang Lee; Gang Huang; Jingnan Shen; Jian Tu.

Ethical approval

Ethical approval for this study was waived by the Institutional Review Board of the First Affiliated Hospital of Sun Yat-sen University since the research involved the analysis of existing, de-identified data from the public SEER database. This study was performed in full compliance with the data use agreements and guidelines stipulated by the SEER program.

Ethical approval for research

N.A.

External funding

Yes.

Funding

J.T. was supported by the National Natural Science Foundation of China (Grant 82472685), Kelin New Star of the First Affiliated Hospital of Sun Yat-sen University, and Youth S&T Talent Support Programme of Guangdong Provincial Association for Science and Technology.

Declaration statement

Jian Tu received funding from: National Natural Science Foundation of China (Grant Number: 82472685), Kelin New Star of the First Affiliated Hospital of Sun Yat-sen University, Youth S&T Talent Support Programme of Guangdong Provincial Association for Science and

Technology.

Acknowledgements

The authors acknowledge financial support from National Natural Science Foundation of China (Grant 82472685), Kelin New Star of the First Affiliated Hospital of Sun Yat-sen University, and Youth S&T Talent Support Programme of Guangdong Provincial Association for Science and Technology and thank the National Cancer Institute (NCI) for providing access to the Surveillance, Epidemiology, and End Results (SEER) database.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejso.2026.111915.

Data availability

The data that support the findings of this study are openly available in the Surveillance, Epidemiology, and End Results (SEER) database at https://seer.cancer.gov/.

References

[1] Ritter J, Bielack SS. Osteosarcoma. Ann Oncol 2010;21(7):320–5. https://doi.org/10.1093/annonc/mdq276. vii.
[2] Rosen G, Murphy ML, Huvos AG, Gutierrez M, Marcove RC. Chemotherapy, en bloc resection, and prosthetic bone replacement in the treatment of osteogenic sarcoma. Cancer 1976;37(1):1–11. https://doi.org/10.1002/1097-0142(197601)37:1<1::aid-cncr2820370102>3.0.co;2-3.
[3] Rosen G, Marcove RC, Caparros B, Nirenberg A, Kosloff C, Huvos AG. Primary osteogenic sarcoma: the rationale for preoperative chemotherapy and delayed surgery. Cancer 1979;43(6):2163–77. https://doi.org/10.1002/1097-0142(197906)43:6<2163::aid-cncr2820430602>3.0.co;2-s.
[4] Picci P, Ferrari S, Bacci G, Gherlinzoni F. Treatment recommendations for osteosarcoma and adult soft tissue sarcomas. Drugs 1994;47(1):82–92. https://doi.org/10.2165/00003495-199447010-00006.
[5] Bacci G, Longhi A, Fagioli F, Briccoli A, Versari M, Picci P. Adjuvant and neoadjuvant chemotherapy for osteosarcoma of the extremities: 27 year experience at Rizzoli institute, Italy. Eur J Cancer (Oxford, England : 1990) 2005;41(18):2836–45. https://doi.org/10.1016/j.ejca.2005.08.026.
[6] Meltzer PS, Helman LJ. New Horizons in the treatment of Osteosarcoma. N Engl J Med 2021;385(22):2066–76. https://doi.org/10.1056/NEJMra2103423.
[7] Biermann JS, Hirbe A, Ahlawat S, et al. Bone cancer, version 2.2025, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw 2025;23(4). https://doi.org/10.6004/jnccn.2025.0017.
[8] Bramwell VH, Burgers M, Sneath R, et al. A comparison of two short intensive adjuvant chemotherapy regimens in operable osteosarcoma of limbs in children and young adults: the first study of the european osteosarcoma intergroup. J Clin Oncol 1992;10(10):1579–91. https://doi.org/10.1200/jco.1992.10.10.1579.
[9] Souhami RL, Craft AW, Van der Eijken JW, et al. Randomised trial of two regimens of chemotherapy in operable osteosarcoma: a study of the European Osteosarcoma Intergroup. Lancet (London, England) 1997;350(9082):911–7. https://doi.org/10.1016/s0140-6736(97)02307-6.
[10] Fuchs N, Bielack SS, Epler D, et al. Long-term results of the co-operative German-Austrian-Swiss osteosarcoma study group's protocol COSS-86 of intensive multidrug chemotherapy and surgery for osteosarcoma of the limbs. Ann Oncol 1998;9(8):893–9. https://doi.org/10.1023/a:1008391103132.
[11] Ferrari S, Smeland S, Mercuri M, et al. Neoadjuvant chemotherapy with high-dose Ifosfamide, high-dose methotrexate, cisplatin, and doxorubicin for patients with localized osteosarcoma of the extremity: a joint study by the Italian and Scandinavian Sarcoma groups. J Clin Oncol 2005;23(34):8845–52. https://doi.org/10.1200/jco.2004.00.5785.
[12] Lewis IJ, Nooij MA, Whelan J, et al. Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma intergroup. J Natl Cancer Inst 2007;99(2):112–28. https://doi.org/10.1093/jnci/djk015.
[13] Le Deley MC, Guinebretière JM, Gentet JC, et al. SFOP OS94: a randomised trial comparing preoperative high-dose methotrexate plus doxorubicin to high-dose methotrexate plus etoposide and ifosfamide in osteosarcoma patients. Eur J Cancer (Oxford, England : 1990) 2007;43(4):752–61. https://doi.org/10.1016/j.ejca.2006.10.023.
[14] Goorin AM, Schwartzentruber DJ, Devidas M, et al. Presurgical chemotherapy compared with immediate surgery and adjuvant chemotherapy for nonmetastatic osteosarcoma: Pediatric oncology group study POG-8651. J Clin Oncol 2003;21(8):1574–80. https://doi.org/10.1200/jco.2003.08.165.
[15] Danese MD, Groundland JS. Effect of chemotherapy and surgery timing on mortality in upper and lower extremity osteosarcoma. J Natl Cancer Inst 2025;117(4):611–8. https://doi.org/10.1093/jnci/djae229.
[16] Xu J, Xie L, Guo W. Neoadjuvant chemotherapy followed by delayed surgery: is it necessary for all patients with nonmetastatic high-grade pelvic Osteosarcoma? Clin Orthop Relat Res 2018;476(11):2177–86. https://doi.org/10.1097/corr.0000000000000387.
[17] Bykov K, Patorno E, D'Andrea E, et al. Prevalence of avoidable and bias-inflicting methodological pitfalls in real-world studies of medication safety and effectiveness. Clin Pharmacol Therapeut 2022;111(1):209–17. https://doi.org/10.1002/cpt.2364.
[18] Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology (Cambridge, Mass) 2008;19(6):766–79. https://doi.org/10.1097/EDE.0b013e3181875e61.
[19] Stensrud MJ, Valberg M, Røysland K, Aalen OO. Exploring selection bias by causal frailty models: the magnitude matters. Epidemiology (Cambridge, Mass) 2017;28(3):379–86. https://doi.org/10.1097/ede.0000000000000621.
[20] Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183(8):758–64. https://doi.org/10.1093/aje/kwv254.
[21] Fu EL. Target trial emulation to improve causal inference from observational data: what, why, and how? J Am Soc Nephrol : JASN 2023;34(8):1305–14. https://doi.org/10.1681/asn.0000000000000152.
[22] Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. JAMA 2022;328(24):2446–7. https://doi.org/10.1001/jama.2022.21383.
[23] Surveillance, epidemiology, and end results (SEER) program (www.seer.cancer.gov)SEER*StatDatabase:Incidence-SEERResearchData,8Registries,Nov2024Sub (1975-2022), National cancer institute, DCCPS, surveillance research program, released April 2025, based on the November 2024 submission. In.
[24] D'Agostino Jr RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17(19):2265–81. https://doi.org/10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b.
[25] Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology (Cambridge, Mass) 2000;11(5):550–60. https://doi.org/10.1097/00001648-200009000-00011.
[26] Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 2000;56(3):779–88. https://doi.org/10.1111/j.0006-341x.2000.00779.x.
[27] VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-Value. Ann Intern Med 2017;167(4):268–74. https://doi.org/10.7326/m16-2607.
[28] Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). JAMA 2002;288(23):2981–97. https://doi.org/10.1001/jama.288.23.2981.
[29] Platt R, Takvorian SU, Septimus E, et al. Cluster randomized trials in comparative effectiveness research: randomizing hospitals to test methods for prevention of healthcare-associated infections. Med Care 2010;48(6 Suppl):S52–7. https://doi.org/10.1097/MLR.0b013e3181dbebcf.
[30] Zhuang R, Xia F, Wang Y, Chen YQ. A surrogate measure for time-varying biomarkers in randomized clinical trials. Mathematics 2022;10(4). https://doi.org/10.3390/math10040584.
[31] Speich B, Taji Heravi A, Schönenberger CM, et al. Nonregistration, discontinuation, and nonpublication of randomized trials: a systematic review. JAMA Netw Open 2025;8(9):e2524440. https://doi.org/10.1001/jamanetworkopen.2025.24440.
[32] Cashin AG, Hansford HJ, Hernán MA, et al. Transparent reporting of observational studies emulating a target trial: the TARGET statement. BMJ (Clinical research ed) 2025;390:e087179. https://doi.org/10.1136/bmj-2025-087179.
[33] Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med 2019;25(10):1601–6. https://doi.org/10.1038/s41591-019-0597-x.
[34] Hernán MA, Dahabreh IJ, Dickerman BA, Swanson SA. The target trial framework for causal inference from observational data: why and when is it helpful? Ann Intern Med 2025;178(3):402–7. https://doi.org/10.7326/annals-24-01871.
[35] Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol 2008;167(4):492–9. https://doi.org/10.1093/aje/kwm324.
[36] Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016;79:70–5. https://doi.org/10.1016/j.jclinepi.2016.04.014.
[37] Didelez V. Commentary: should the analysis of observational data always be preceded by specifying a target experimental trial? Int J Epidemiol 2016;45(6):2049–51. https://doi.org/10.1093/ije/dyw032.
[38] Fu EL, Evans M, Carrero JJ, et al. Timing of dialysis initiation to reduce mortality and cardiovascular events in advanced chronic kidney disease: nationwide cohort study. BMJ (Clinical research ed) 2021;375:e066306. https://doi.org/10.1136/bmj-2021-066306.
[39] Alaseem A, Alanezi M, Almehrij F, et al. Comparative outcomes of neoadjuvant chemotherapy versus upfront surgical resection in osteosarcoma: a systematic review and meta-analysis. World J Surg Oncol 2025;23(1):446. https://doi.org/10.1186/s12957-025-04115-3.

[40] Frei E. 3rd. Curative cancer chemotherapy. Cancer Res 1985;45(12 Pt 1):6523–37.
[41] Aboulafia AJ, Malawer MM. Surgical management of pelvic and extremity osteosarcoma. Cancer 1993;71(10 Suppl):3358–66. https://doi.org/10.1002/1097-0142(19930515)71:10+<3358::aid-cncr2820711738>3.0.co;2-o.
[42] Bielack SS, Kempf-Bielack B, Delling G, et al. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol 2002;20(3):776–90. https://doi.org/10.1200/jco.2002.20.3.776.
[43] Bacci G, Bertoni F, Longhi A, et al. Neoadjuvant chemotherapy for high-grade central osteosarcoma of the extremity. Histologic response to preoperative chemotherapy correlates with histologic subtype of the tumor. Cancer 2003;97(12):3068–75. https://doi.org/10.1002/cncr.11456.
[44] Bacci G, Picci P, Ruggieri P, et al. Primary chemotherapy and delayed surgery (neoadjuvant chemotherapy) for osteosarcoma of the extremities. The istituto Rizzoli experience in 127 patients treated preoperatively with intravenous methotrexate (high versus moderate doses) and intraarterial cisplatin. Cancer 1990;65(11):2539–53. https://doi.org/10.1002/1097-0142(19900601)65:11<2539::aid-cncr2820651125>3.0.co;2-m.
[45] Bacci G, Ferrari S, Tienghi A, et al. A comparison of methods of loco-regional chemotherapy combined with systemic chemotherapy as neo-adjuvant treatment of osteosarcoma of the extremity. Eur J Surg Oncol : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2001;27(1):98–104. https://doi.org/10.1053/ejso.2000.1056.
[46] Meyers PA, Schwartz CL, Krailo MD, et al. Osteosarcoma: the addition of muramyl tripeptide to chemotherapy improves overall survival–a report from the children's oncology group. J Clin Oncol 2008;26(4):633–8. https://doi.org/10.1200/jco.2008.14.0095.
[47] Marina NM, Smeland S, Bielack SS, et al. Comparison of MAPIE versus MAP in patients with a poor response to preoperative chemotherapy for newly diagnosed high-grade osteosarcoma (EURAMOS-1): an open-label, international, randomised controlled trial. Lancet Oncol 2016;17(10):1396–408. https://doi.org/10.1016/s1470-2045(16)30214-5.
[48] Piperno-Neumann S, Le Deley MC, Rédini F, et al. Zoledronate in combination with chemotherapy and surgery to treat osteosarcoma (OS2006): a randomised, multicentre, open-label, phase 3 trial. Lancet Oncol 2016;17(8):1070–80. https://doi.org/10.1016/s1470-2045(16)30096-1.
[49] Gaspar N, Occean BV, Pacquement H, et al. Results of methotrexate-etoposide-ifosfamide based regimen (M-EI) in osteosarcoma patients included in the French OS2006/sarcome-09 study. Eur J Cancer (Oxford, England : 1990) 2018;88:57–66. https://doi.org/10.1016/j.ejca.2017.09.036.